

seq_name: pir2:152615

seq_documentation_block:
 gene Rhd protein - human
 C:Species: Homo sapiens (man)
 C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 05-Nov-1999
 C:Accession: 152615
 R: Huang, C. H.; Reid, M. E.; Chen, Y.
 R: Blood 86, 784-790, 1995
 A: Title: Identification of a partial internal deletion in the Rh locus causing the human
 A: Reference number: 152615; MUID:95329738
 A: Accession: 152615
 A: Status: preliminary; translated from GB/EMBL/DBJ
 A: Molecule type: mRNA
 A: Residues: 1-417 <RES>
 A: Cross-references: GB:S78509; NID:9999309; PIDN:AAB34852.1; PID:9999310
 C: Genetics:
 Gene: Rhd

alignment_scores:
 Quality: 2100.00 Length: 417
 Ratio: 5.060 Gaps: 0
 Percent Similarity: 99.520 Percent Identity: 99.281

alignment_block:
 US-09-600-714-41 x 152615 ..

Align seg 1/1 to: 152615 from: 1 to: 417

```

1  ATGAGCTCTAGTACCGCGGTCTGTCGGCGGCGCTGACCCCTGTGGGC 50
|||||
1  MetSerSerIyIyrProAlrGSerValAlrGAlrGylSerProIleuGyl 17
17  AlenThrIeuGlunAlaIleuIleuIleuPheThrIleuHIST 34
|||||
51  CCTAACACTGAGACAGCTCTCATCTCTCTCTATTTTATACCACT 100
|||||
101  ATGAGCTCTCTTAGAGGATCAAAAGGGCTGCGGATCTCATTAATT 150
|||||
34  YrAspAlaSerIeuGlunAspGlnIyIeuValAlaSerIyIeuVal 50
|||||
151  GGCAAGATGACCGGTGATGGCGGCAATGGCTTGGCTTCCGACCTC 200
|||||
51  GlnIeuAspIeuThrValMetAlaIleuGlnIyIeuGlnPheIeuHIS 67
|||||
201  GAGTTTCCGAGACACAGCTGAGAGAGTGTGCTTCAACCTCTTCATGC 250
|||||
67  rSerPheIeuAlrGAlrHisSerIerPserIerValAlaIleuAsnIeuPheMetL 84
|||||
251  TGGCGCTTGTGTGCGAGTGGCAATCTCTGCTGAGCGCTTCTGAGCCAG 300
|||||
84  euAlaIeuGlnIyAlGlnIeuPheAlaIleuIleuAspGlnPheIeuSerGln 100
|||||
301  TTCCCTTGTGGAAGTGTGATCATCACCTGATTCAGTATTCGGCTGGCCAC 350
|||||
101  PheProSerGlnIyValIleuThrIeuPheSerIleuIleuGlnAlaH 117
|||||
351  CATGAGTCTTGTGGTGTGATCTCATGATGATCTGTCTTGGGGAAG 400
|||||
117  rMetSerAlaIleuSerValIleuIleuSerValAlaIleuGlnIyVal 134
|||||
401  TCAACTTGGCGAGTGTGTGATGTGTGCTGTGAGAGTGAACCTTTA 450
|||||
134  AlAsnIleuAlaIleuValIleuValMetValIleuValGlnValThrAlaIeu 150
|||||
451  GGCACCTCGAGAGATGCTATCATATATCTTCAACAGACTACACAT 500
|||||
151  GlnAlaIeuAlrMetValIleuSerAlnIleuPheAsnIleuAspIyrHisHe 167
|||||
501  GAACATGATGACATCTACGTGTGTGCGACGATATTTTGGCTGTCTGTG 550
|||||

```

seq_name: pir2:137005

seq_documentation_block:
 Rhesus-like protein - chimpanzee
 C:Species: Pan troglodytes (chimpanzee)
 C:Date: 04-Oct-1996 #sequence_revision 04-Oct-1996 #text_change 21-Jul-2000
 C:Accession: 137005
 R: Salviagnol, I.; Blancher, A.; Calvas, P.; Clayton, J.; Socha, W.W.; Collin, Y.; Ruffi
 Biochem. Genet. 32, 201-221, 1994
 A: Title: Molecular genetics of chimpanzee Rh-related genes: their relationship with t
 A: Reference number: 137003; MUID:95085595

```

167  tAsnMetMetHisIleIeuThrMetPheAlaAlaIyrPheGlnIeuSerValA 184
|||||
551  CCTGTGCTGTGCCAAAGCCTCTACCCGAGGGAACGAGATTAAGATCAG 600
|||||
184  lArIryCysIeuProIySProIeuProGlnIyThrGlnAspIySAspGln 200
|||||
601  ACAGCAACGATACCCAGATTGTCTGCCATGCTGGGGCCCTCTTCTGTG 650
|||||
201  ThrAlaThrIleProSerIeuSerAlaMetIeuGlnIyAlaIeuPheIeuTr 217
|||||
651  GATCTTGTGGCAAGTTTCAACTCTGCTGCTGAGAGTCAATCGAAA 700
|||||
217  pMetPheIerProSerPheAsnSerAlaIleuAlaIeuAlrSerProIleuH 234
|||||
701  GGAAGAAATGCCGTGTTCAACACTACTAGCTGTAGCAGTCAGCTGTG 750
|||||
234  rGlnAsnAlaIleuAlaPheAsnThrIyThrAlaValAlaIleuSerVal 250
|||||
751  ACAGCATCTCAGGGTCACTCTTGTGCTACCCCAAGGAAGATCAGCAA 800
|||||
251  ThrAlaIleuSerGlnSerIeuAlaIleuAlaIleuProGlnIyIleuSerIy 267
|||||
801  GACTATGTGCACAGTCCGGTGTGGCAGAGAGCGTGGCTGTGGTACCT 850
|||||
267  sThrIyValHisSerAlaValIleuAlaGlnIyGlnIyAlaValAlcIyHrs 284
|||||
851  CGTGTCACTGATCCCTTCTCCGTGCTTGCATGGCTGAGTGGCTGTG 900
|||||
284  erCysHisIleuIleuProSerProIleuAlaMetValIleuGlnIyVal 300
|||||
901  GCTGGGCTGATCTCCGTGCGGGGAGCCAGTACTGCGGGGTTGTGA 950
|||||
301  AlaGlnIleuIleuSerValGlnIyAlaIyIeuProGlnIyCysCysAs 317
|||||
951  CCGAGTGTGGGGATGCCACAGCTCATCATGAGGCTGACAACTTCACT 1000
|||||
317  nArgValIleuGlnIyIleuProIleuSerIleuMetGlnIyIyrAsnIleuSerL 334
|||||
1001  TGCTGGGTCTCTGAGATGATCATCTACATTTGCTGCTGCTGTGAT 1050
|||||
334  euIeuGlnIleuIleuGlnIyIleuIleuIleuIleuIleuValIleuAsp 350
|||||
1051  ACCGTGAGCCCGGCAATGATGATTTGGCTTCCAGGTCCTTCAGAT 1100
|||||
351  ThrValGlnAlaGlnIyAsnGlnIyMetIleuGlnIyPheGlnIleuIleuSerI 367
|||||
1101  TGGGAACTCACTTGGCCATCGTATGATGCTCACGCTGCTGCTCGA 1150
|||||
367  eGlnIleuIleuSerIeuAlaIleuAlaIleuAlaIleuThrIleuSerIleuIleuTr 384
|||||
1151  CAGTTTCTCTTAATCTTAATATATGAAAGCACTCATGAGGCTTAA 1200
|||||
384  hGlnIleuIleuIleuAsnIleuIleuIleuIleuIleuIleuIleuIleu 400
|||||
1201  TATTTTGTATGACCAAGTTTCTGGAAGTTTCTCATTTGGCTGTGATT 1250
|||||
401  TyrPheAspAspGlnValPheThrIySPherProHisIleuAlaValGln 417
|||||
1251  T 1251
417 e 417

```

THIS PAGE BLANK (USPTO)

A:Accession: I37005
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-417 <RES>
 A:Cross-references: GB:L37050; NID:9606991; PIDN:AAA65624.1; PID:9606992

alignment_scores:
 Quality: 1975.00 Length: 417
 Ratio: 4.889 Gaps: 0
 Percent Similarity: 96.882 Percent Identity: 93.046

alignment_block:
 US-09-600-714-41 x I37005

Align seg 1/1 to: I37005 from: 1 to: 417

```

1 ATGAGCTCTAAGTACCCGGGCTGTGCTCCGGGCTGCTGCCCTCTGGGC 50
|||||
1 MetSerSerLysTyrProArgSerValArgArgCysLeuProLeuCysAl 17
51 CCTAACCTGGAAGCAGCTCATCTCTCTCTATTTTATTTTACCACT 100
|||||
17 AleuThrLeuGluAlaAlaLeuIleLeuPheTyrPhePheThrGlnT 34
101 ATGACGCTTCCTTAGAGATCAAAAGGGGCTCGGCACTCCTATCAAGT 150
|||||
34 TyrAspAlaSerLeuGluAspGlnGlyLeuValAlaSerTyrGlnVal 50
151 GGCCTAATATGACCGGTGATGAGCGGCTGCTGGGCTTCCACCTC 200
|||||
51 GlyGlnAspLeuThrValMetAlaAlaIleGlyPheGlyPheLeuThrSe 67
201 GAGTTCCTGGAGACACAGCTGAGCAGTGTGCTTCAACCTCTTCATGC 250
|||||
67 rSerPheArgArgHisSerTyrPheSerSerValAlaPheSerLeuPheMetL 84
251 TGGCGCTTGGTGTGCAAGTGGGCAATCCTGCTGAGCGGCTTCTGAGCG 300
|||||
84 euAlaLeuGlyValGlnTPrAlaIleLeuLeuAspGlyPheLeuSerGln 100
301 TTCCCTCTGGGAAGGTGCTCATACACACTGTTCAGTATCGGCTGGCCAC 350
|||||
101 PheProGlyGlyValValIleThrLeuPheSerIleAlaGlyLeuAlaThr 117
351 CATTGAGCTTGTGCGGTGATCTCAGTGCATGATGCTGTCTTGGGAAG 400
|||||
117 rThrSerAlaLeuSerValIleLeuIleSerValAlaSerValLeuGlyLysV 134
401 TCAACTGGGCGCAGTGTGCTGATGCTGTGCTGAGAGTGACAGCTTTA 450
|||||
134 alaAsnLeuValGlnLeuValValMetValLeuValGlnValThrAlaLeu 150
451 GGCACCTGAGGATGCTCATCATATATCTTCAACACACTCCACAT 500
|||||
151 GlyThrValArgMetValIleSerAsnIlePheAsnThrAspTyrHisMe 167
501 GAACATATGACATCTACGCTGCGAGCCTATTTGGGCTGCTGTGG 550
|||||
167 tAsnLeuMetHisIleTyrValIleAlaIleAlaTyrPheGlyLeuSerValA 184
551 CCTGGTCCCTCCAAAGCCTCTTACCGAGGGAAGGAGATAAAGATCAG 600
|||||
184 latrPcysLeuProLysProLeuProLysGlyThrGluAspGlyAspGln 200
601 ACAGACAGATACCCAGTTGTCTGCAATGCTGGGCGCCCTTCTTGTG 650
|||||
201 IleAlaThrIleProSerLeuSerAlaMetLeuGlyAlaLeuPheLeuTr 217
651 GATGTGCTGGCAGTTTCAACTGCTGCTGCTGAGAAATCCAAATCGAAA 700
|||||
217 pMetPheTrpProSerPheAsnSerAlaIleLeuLeuArgSerProIleGluA 234

```

```

701 GGAAGAAATGCCGTTCACACACTACTATGCTGTACGACGACCTGTG 750
|||||
234 rglYssAsnAlaValPheAsnThrTyrTyrAlaValAlaValSerVal 250
751 ACAGCATTTCAGGATCATCTGCTGCTGCTGCTGCTGCTGCTGCTG 800
|||||
251 ThrAlaIleSerIleSerLeuAlaHisProGlnGlyLysIleSerMe 267
801 GACTATATGTCACAGTCCGCTGTGCGAGAGGCGTGGCTGTGGTACCT 850
|||||
267 tSerTyrMetHisAsnAlaValLeuAlaGlyGlyValAlaValGlyThrS 284
851 CGTGTACACTGATCCCTTCTCCGCTGCTGCTGCTGCTGCTGCTGCTG 900
|||||
284 erCysHisLeuIleProSerProTyrPheAlaMetValLeuGlyLeuVal 300
901 GCTGGGCTGATCTCCGCTGGGAGGCAAGTACTGCTGCGGGGTGTGTA 950
|||||
301 AlaGlyLeuIleSerValGlyGlyAlaLysTyrLeuProGlyCysAs 317
951 CCGAGTCTGGGATGCCACAGCTCCATCATGAGGCTACACTGACGT 1000
|||||
317 narGValLeuGlyIleProHisSerSerValMetGlySerAsnPheSer 334
1001 TGCTGGCTGCTTGGAGATCATCTACATTTGCTGCTGCTGCTGCTGAT 1050
|||||
334 rPheGlyLeuLeuGlyGlnIleTyrIleValLeuValAlaArgHis 350
1051 ACCGTGCGACCGGCAATGCGATGATGCTTCCAGTCCCTCCAGCAT 1100
|||||
351 ThrIleTrpAsnGlyAsnGlyMetIleGlyPheGlnAlaLeuLeuArgI 367
1101 TGGGAACTACGCTTGGCCATCGTATGCTCTACGCTGCTGCTCTGTA 1150
|||||
367 eglYgluPheSerLeuAlaThrThrIleAlaLeuThrSerGlyLeuLeuThr 384
1151 CAGGTGCTCTCTTAATCTTAAATATGAAAGACACTCATGAGCTAAA 1200
|||||
384 hrGlyLeuLeuLeuAsnLeuLysIleTrpLysAlaProHisGlnAlaLys 400
1201 TATTTGATGACCAAGTTTCTGGAAGTTCTCATTTGCTGTGGATT 1250
|||||
401 TyrPheAspAspGlnValPheTrpLysPheProHisLeuAlaValGlyPh 417
1251 T 1251
417 e 417

```

seq_name: p1r2:I37076

seq_documentation_block:

Rhesus-like protein - gorilla
 C:Species: gorilla (gorilla)
 C:Date: 04-Oct-1996 #sequence revision 04-Oct-1996 #text change 21-Jul-2000
 C:Accession: I37076
 R:Salvignol, I.; Blancher, A.; Calvas, P.; Clayton, J.; Socha, W.W.; Collin, Y.; Ruffi
 Biochem. Genet. 32, 201-221, 1994
 A:Title: Molecular genetics of chimpanzee Rh-related genes: their relationship with t
 A:Reference number: I37003; MUID:95085595
 A:Accession: I37076
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-417 <RES>
 A:Cross-references: GB:L37053; NID:9607011; PIDN:AAA65627.1; PID:9607012

alignment_scores:
 Quality: 1953.00 Length: 417
 Ratio: 4.870 Gaps: 0
 Percent Similarity: 96.163 Percent Identity: 92.806

alignment_block:
 US-09-600-714-41 x I37076

THIS PAGE BLANK (USPTO)

Align seg 1/1 to: 137076 from: 1 to: 417

```

1 ATGAGCTCTAAGTACCCGGGTCTGTCCGGCGCTGCTCCCTTGAGC 50
1 MetSerSerLysTyrProArgSerValAlaIgcysLeuProLeuLysAla 17
51 CCTAACACTGGAAGAGCTCTCATCTCTCTCTCTATTTTATCCACT 100
17 aleuThrLeuGlnAlaAlaLeuThrLeuPheThrPhePheThrHisT 34
101 ATGAGCTCTCTTAGAGATCAAAAGGGCTGCTGAGCATCTATCAAGT 150
34 yrasPalSerLeuGlnaspGlnLysGlyLeuValAlaSerTyrGlnVal 50
151 GGCCAGATCTGACCGTGTATGGCGGCATTTGGCTGGCTCTCACTC 200
51 GlnGlnAspLeuThrValMetAlaAlaIleGlyPheGlyPheLeuThr 67
201 GAGTTTCCGGAACACAGCTGGAGAGTGGCTTCAACCTCTTCATGC 250
67 rSerPheArgGlyHisSerTyrSerSerValAlaPheAsnLeuPheMet 84
251 TGGCGCTTGCTGTGCTGCGGCAATCTGCTGAGCGCTTCTGAGCAG 300
84 euAlaLeuGlyValGlnTrrPalaIleLeuLeuaspGlyPheLeuSerGln 100
301 TTCCCTCTTGGGAAGGTGTCATCACACTTTCAGTATTCGGCTGGCC 350
101 PheProGlyLysValValIleThrLeuPheSerIleArgLeuAlaThr 117
351 CATGAGTCTTGTGCTGCTGATCTCAGTGATGCTGTCTTGGGAGAG 400
117 rMetSerAlaLeuSerValLeuIleSerAlaGlyAlaValLeuGlyTyr 134
401 TCAACTGGCGCAGTTGGTGTGATGCTGTGCTGAGCTGAGCAGCTT 450
134 alsnLeuValGlnLeuValValMetValLeuValGlnValIleThrAlaLeu 150
451 GGCACTGAGAGTGTCTCATCAATATCTTCAACACAGACTACACAT 500
151 GlyThrMetArgMetValIleSerAsnIlePheAsnThrAspTyrHisMe 167
501 GAACATGATGACATCTACTGCTGTGCGAGCCTATTTGGCTGCTGCTG 550
167 tAsnMetMetHisIleTyrValPheAlaAlaCysPheGlyLeuSerVal 184
551 CCGTGTGCTGCGCAAGCTCTACCCGAGGAGGAGGAGATTAAGATCAG 600
184 latrPcysLeuProLysProLeuAlaLysGlyThrGlnaspLysAspGln 200
601 ACAGCAAGATACCAAGTTTGTCTGCCATGCTGGCGCCTCTTCTGTG 650
201 ThrIleThrIleProSerLeuSerAlaMetLeuGlyAlaLeuPheLeuTr 217
651 GATGTTCTGCCAAGTTTCAACTGCTGCTGCTGAGAAGTCCAAATCGAA 700
217 pMetPheTrpSerPheAsnSerAlaLeuLeuArgSerProIleGlnAla 234
701 GGAGAAATCGCGTGTCAACACTACTATGCTGTAGCAGTACAGCGTGTG 750
234 rGlyAsnAlaValAlaPheAsnThrTyrTrrAlaValAlaValSerValVal 250
751 ACAGCATCTCAGGATCATCTTGCTGCTACCCCAAGGAAAGATCAGCAA 800
251 ThrAlaIleSerValSerSerLeuAlaHisProGlnGlyLysIleAsnMe 267
801 GACTTATGTGACAGTGGCGGTGTGGCAGAGAGCGTGGCTGTGGSTACT 850
267 tThrTyrMetHisAsnAlaValLeuAlaGlyGlyValAlaValAlaThrS 284
851 CGTGCACCTGATCCCTTCTCCGCGCTTGCATGCTGCTGGCTGTGTG 900

```

```

284 erCysHisLeuIleProSerProTrpLeuAlaMetValLeuGlyLeuVal 300
901 GCTGGCTGATCTCCGTGCGGAGGCCAAGTACCTGCCGGGTGTGTAA 950
301 AlaGlyLeuIleSerIleGlyGlyAlaLysCysLeuProGlyCysCysAs 317
951 CCGAGTGTGGGGATTTCCACAGCTCCATCATGGCTACCAACTGACCT 1000
317 nArgValLeuGlyIleHisAspSerSerValMetHisTyrAsnPheSer 334
1001 TGCCTGGGTGCTGTGAGAGATCATCTACATTTGCTCTGCTGCTGTAT 1050
334 euLeuGlyLeuLeuGlyGlnIleTyrIleValLeuLeuValLeuasp 350
1051 ACCGTGCGAGCCGCAATGGCATGATTGCTTCCAGTCCCTCTCAGCAT 1100
351 ThrValGlyAlaGlyAsnGlyMetValGlyPheGlnValIleValSerI 367
1101 TGGGGAATCAGCTTGGCCATGCTGATACCTCTCAGCTGTGCTCTGA 1150
367 eGlyGlnLeuSerLeuAlaIleValIleAlaLeuThrSerGlyLeuLeuT 384
1151 CAGGTTGCTCTTAATCTTAATATATGGAAGCACTCATAGGCTATA 1200
384 hGlyLeuLeuLeuAsnLeuLysIleTrrPysAlaProHisAlaAlaLys 400
1201 TATTTGATGACCAAGTTTTCGAAGTTTCCATTTGCTGCTGTGATT 1250
401 TyrPheAspAspGlnValIleThrTrrPysPheProHisLeuAlaValGly 417
1251 T 1251
417 e 417

```

seq_name: p1r2:137004

```

seq_documentation_block:
Rhesus-like protein - chimpanzee
C:Species: Pan troglodytes (chimpanzee)
C:Date: 04-Oct-1996 #sequence_revision 04-Oct-1996 #text_change 21-Jul-2000
C:Accession: 137004
R:Salvignol, I.; Blancher, A.; Calvas, P.; Clayton, D.; Socha, W.W.; Collin, Y.; Ruffi
Blochem. Genet. 32, 201-221, 1994
A:title: Molecular genetics of chimpanzee Rh-related genes: their relationship with t
A:Reference number: 137003; MUID:95085595
A:Accession: 137004
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-417 <RES>
A:Cross-references: GB:137049; NID:9606989; PIDN:AA65623.1; PID:9606990

```

alignment_scores:

Quality	Length
1952.00	417
Ratio: 4.868	Gaps: 0
Percent Similarity: 96.163	Percent Identity: 92.566

alignment_block:

us-09-600-714-41 x 137004

Align seg 1/1 to: 137004 from: 1 to: 417

```

1 ATGAGCTCTAAGTACCCGGGTCTGTCCGGCGCTGCTCCCTTGAGC 50
1 MetSerSerLysTyrProArgSerValAlaIgcysLeuProLeuLysAla 17
51 CCTAACACTGGAAGAGCTCTCATCTCTCTCTATTTTATCCACT 100
17 aleuThrLeuGlnAlaAlaLeuThrLeuPheThrPhePheThrHisT 34
101 ATGAGCTCTCTTAGAGATCAAAAGGGCTGCTGAGCATCTATCAAGT 150
34 yrasPalSerLeuGlnaspGlnLysGlyLeuValAlaSerTyrGlnVal 50

```

THIS PAGE BLANK (USPTO)

151 GGCACAGATCTGACCGGTGATGGCGCCATTGGCTTGGGCTTCCACCTC 200
 |||||
 51 GlyGlnaspLeuthrValMetAlaAlaIleGlyPheGlyPheLeuthrSe 67
 |||||
 201 GAGTTTCCGAGACACAGCTGAGACAGTGTGGCTTCAACCTTTCATGC 250
 |||||
 67 rSerPheArgArgHisSerTrpSerSerValAlaPheSerLeuPheMetL 84
 |||||
 251 TGGCGCTTGGTGGCAGTGGGCAATCCTGCTGACGCGCTTCCGAGCCAG 300
 |||||
 84 euAlaLeuGlyValGlnTrpAlaIleLeuLeuaspGlyPheLeuSerGln 100
 |||||
 301 TTCCCTTCCGAGAGGTGTCATCACACACTGTCAGTATTCGGCTGGCCAC 350
 |||||
 101 PheProProGlyValValIleThrLeuPheSerIleArgLeuAlaThr 117
 |||||
 351 CAGTATGCTTGTGCTGCTGATCTCAGTATGCTGTCTTGGGGAAG 400
 |||||
 117 rThrSerAlaLeuSerValLeuIleSerValAspAlaValLeuGlyLysV 134
 |||||
 401 TCACTTGGCGCAGTTGGTGTGATGGTGGTGGGAGGTGACAGCTTTA 450
 |||||
 134 AlasnLeuValGlnLeuValValMetValGluValIleThrAlaLeu 150
 |||||
 451 GGCACCTGAGATGTGTCATCATATATCTTCAACACAGACTACACACT 500
 |||||
 151 GlyThrValArgMetValIleSerAsnIlePheasnThrAspTrpHisMe 167
 |||||
 501 GAACATGATGCACACTGATGCTTGGCAGCCATATTTGGGCTGTCTGG 550
 |||||
 167 tAsnLeuMetHisIleLeuValPheAlaAlaLysPheGlyLeuSerValA 184
 |||||
 551 CCGTGGCTGCCAAGCCTCTACCCGAGGGAACGGAGATTAAGATCAG 600
 |||||
 184 lArpCysLeuProLysProLeuProLysGlnGluAspPylAspGln 200
 |||||
 601 ACAACAAGATACCACTTGTGTCATGCTGCGCGCTCTCTCTCTG 650
 |||||
 201 IleIleThrIleProSerLeuSerAlaMetLeuGlyValAlaLeuPheLeuTr 217
 |||||
 651 GATGCTTGGCCAAAGTTTCAACTCTGCTGTGAGAAAGTCCAAATCGAAA 700
 |||||
 217 pMetPheTrpProSerPheasnSerAlaLeuLeuArgSerProIleGluA 234
 |||||
 701 GGAAGATGCCGTGTACACACTCATGCTGTAGAGACAGCAGCTGTG 750
 |||||
 234 rGlySasnAlaValPheasnThrLysTrpAlaValAlaValSerValVal 250
 |||||
 751 ACAAGCATCTCAGGGTTCATCTGCTGCTCACCACCCCAAGAGATCAGCAA 800
 |||||
 251 ThrAlaIleSerGlySerSerIleAlaHisProGlnGlyLysIleSerMe 267
 |||||
 801 GACTTATGTGCACAGTCGGTGTGGCAGAGCGGTGCTGGTGA 850
 |||||
 267 tSerTyrMetHisasnAlaValLeuAlaGlyLysValAlaValIlyThrS 284
 |||||
 851 CGTGTACACGTGATCCCTTCCCGGCTGCGCATGAGTGGCGGCTGTG 900
 |||||
 284 eCySHisLeuIleLysSerProTrpLeuAlaMetValLeuGlyLeuVal 300
 |||||
 901 GCTGGCTGATATCCGTCGGGAGACCAAGTACCTGCCGGGCTTTGTA 950
 |||||
 301 AlaGlyLeuIleSerIleGlyLysAlaLysTrpLeuProGlyCysCysAs 317
 |||||
 951 CCGAGTGTGGGGATTCGCCACAGCTCCATCATGGCTACAACTTCAGCT 1000
 |||||
 317 nArgValLeuGlnGlyLysTrpHisSerSerValMetHisTyrAsnPheSerL 334
 |||||
 1001 TGTGGGCTGCTCTGGAGAGATCATCATGATGGTGGCTGCTGAT 1050
 |||||
 334 eLeuGlyLeuLeuGlyLysIleLeuTrpIleValLeuLeuValHisHis 350

1051 ACCGTGGAGCCGGCAATGATGATGGCTTCCAGCTCCGACGAT 1100
 |||||
 351 ThrValTrpAsnGlySasnGlyMetIleGlyPheGlnValAlaLeuLysGln 367
 |||||
 1101 TGGGAGACTCAGCTTGGCCATGTGATAGTCTCAAGTGTGGCTTCCCTAA 1150
 |||||
 367 eGlyGlnPheSerLeuAlaThrThrIleAlaLeuThrSerGlyLeuLeuTr 384
 |||||
 1151 CAGCTTGGCTCCAAATCTTAAATATGGAAAGCACCTCATGAGGCTAA 1200
 |||||
 384 hGlyLeuLeuLeuSasnLeuLysIleTrpLysAlaProHisAlaAlaLys 400
 |||||
 1201 TATTTGATGACCAAGTGTGTGGAAGTTTCCATTTGGCTGTGGAT 1250
 |||||
 401 TyrPheAspAspAlaValAlaPheTrpLysPheProHisLeuAlaValGln 417
 |||||
 1251 T 1251
 417 e 417

seq_name: p1r2:A30405

seq_documentation_block:

erythrocyte membrane protein Rh (unknown specificity) - human

N:Alternate names: blood group Rh antigen

C:Species: Homo sapiens (man)

C:Date: 05-Oct-1990 #sequence revision 05-Oct-1990 #text change 21-Jul-2000

C:Accession: A30405; S13060; S154273; S40515; S40516; S40517; S02085; S02087; A32509;

R:Cherif-Zahar, B.; Bloy, C.; Le Van Kim, C.; Blanchard, D.; Bailly, P.; Hermand, P.;

Proc. Natl. Acad. Sci. U.S.A. 87, 6243-6247, 1990

A:Title: Molecular cloning and protein structure of a human blood group Rh polypeptid

A:Reference number: A30405; MUID:90349591

A:Accession: A30405

A:Molecule type: mRNA

A:Residues: 1-417 <CHR>

A:Cross-references: GB:M44015; NID:g337398; PIDN:AAA36567.1; PID:g337399

R:Avant, N.D.; Ridgwell, K.; Tanner, M.J.A.; Anstee, D.J.

Biochem. J. 271, 821-825, 1990

A:Title: cDNA cloning of a 30 kDa erythrocyte membrane protein associated with Rh (Rh

A:Reference number: S13060; MUID:91058522

A:Accession: S13060

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-417 <RES>

A:Cross-references: EMBL:X54534; NID:g36017; PIDN:CAA38401.1; PID:g36018

R:Kajli, E.; Umenishi, F.; Iwamoto, S.; Ikemoto, S.

Hum. Genet. 91, 157-162, 1993

A:Title: Isolation of a new cDNA clone encoding an Rh polypeptide associated with the

A:Reference number: S14273

A:Accession: S14273

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-417 <RES>

A:Cross-references: GB:S57967; NID:g299049; PIDN:AA26080.1; PID:g299050

R:Moumou, I.; Collin, Y.; Cherif-Zahar, B.; Carttron, J.P.; Le Van Kim, C.

Nature Genet. 5, 62-65, 1993

A:Title: Molecular genetic basis of the human Rhesus blood group system.

A:Reference number: S40515; MUID:94035121

A:Accession: S40515

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-225, 'A', 227-301, 'F', 303-417 <MOU1>

A:Accession: S40516

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-225, 'A', 227-301, 'F', 303-417 <MOU2>

A:Accession: S40517

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-15, 'C', 17-59, 'I', 61-67, 'S', 69-102, 'S', 104-225, 'A', 227-301, 'F', 303-417 <

Biochem. J. 256, 1043-1046, 1988

A:Title: Protein-sequence studies on Rh-related polypeptides suggest the presence of

THIS PAGE BLANK (USPTO)

A:Reference number: S02085; MUID:89134163
 A:Accession: S02085
 A:Molecule type: protein
 A:Residues: 2-11,'X',13-33 <AVE2>
 A:Note: this peptide was designated blood group Rh-related protein D30
 A:Accession: S02087
 A:Molecule type: protein
 A:Residues: 2-11,'X',13-14 <AVE3>
 A:Note: this peptide was designated blood group Rh-related protein R6A32
 R:Bløy, C.; Blanchard, D.; Dahr, W.; Beyreuther, K.; Salmon, C.; Cartton, J.P.
 Blood 72, 661-666, 1988
 A:title: Determination of the N-terminal sequence of human red cell Rh(D) polypeptide an
 A:Reference number: A32509; MUID:88294325
 A:Accession: A32509
 A:Molecule type: protein
 A:Residues: 2-11,'L',13-15,'X',17 <BL0>
 R:Shoorl, A.M.; Smith, B.L.; Agre, P.
 Proc. Natl. Acad. Sci. U.S.A. 85, 4042-4045, 1988
 A:title: Polymorphism in the M-32,000 Rh protein purified from Rh(D)-positive and -neg
 A:Reference number: A30216; MUID:88234555
 A:Accession: A30216
 A:Molecule type: protein
 A:Residues: 2-11,'X',13-21 <SAB>
 R:Suyma, K.; Goldstein, J.; Aebersold, R.; Kent, S.
 Blood 77, 411, 1991
 A:title: Regarding the size of Rh proteins.
 A:Reference number: A61174; MUID:91091537
 A:Accession: A61174
 A:Molecule type: protein
 A:Residues: 401-407,'DI' <SU>
 C:Keywords: alternative splicing; erythrocyte; surface antigen; transmembrane protein

alignment_scores:
 Quality: 1934.00 Length: 417
 Ratio: 4.872 Gaps: 0
 Percent Similarity: 95.204 Percent Identity: 91.607

alignment_block:
 US-09-600-714-41 x A30405 ..

Align seg 1/1 to: A30405 from: 1 to: 417

```

1 ATAGAGCTTAAGTACCGCGGCTGTCTGCGGCGCTGCGCCCTCTGGG 50
  |||||||
1 MetSerSerIysTyrProArgSerValArgArgCysLeuProLeuIrrp 17
51 CTTAAAGCTGGAAGCAGCTCATCTCTCTCTCTATTTTACCACAT 100
  |||||||
17 aleuthleuglunlaialaleuileleuleuPheTyrPhehethrhnst 34
101 ATGACGCTTCTTATGAGATCAAAAGGCGCTGTCGATCCATCAAGTT 150
  |||||||
34 yraspalaserleuglualspolnlysglyleuValAlaserTyrGlnVal 50
151 GGGCAAGATCTGACCGTGATGGCGCCATTTGGCTGGCTCCACACTC 200
  |||||||
51 GlyIlnaspLeuThrValMetAlaIaleuenglyleuGlyPheleuThnSe 67
201 GAGTTCGCGAGACACAGCTGAGCAGTGTGGCTTCAACCTTTCATGC 250
  |||||||
67 TAsnPhenArgHisSerTrrPserSerValAlaIahPheasnleuPheMetL 84
251 TGGCGCTTGTGTGAGTGGCAATCTGTGTGGAGCGCTTCTGAGCCAG 300
  |||||||
84 eualaleuGlyValGlnIrrpAlaIleleuLeuaspGlyPheleuSerGln 100
301 TTTCCTTTGGAGAGTGTGTATCATCACTGTTCATTTGGCTGGCCAC 350
  |||||||
101 PheProPogIlyValValIleThrleuPheSerIleArgLeuAlaTh 117
351 CATGAGTGTGTGTGGTGTGATCTCAAGTGGATCTGTCTTGGGAGAG 400
  |||||||

```

```

117 rMetSerAlaMetSerValleuIleSerAlaGlyAlaValleuGlyLysV 134
401 TCAACTTGGCGCAGTTGGTGTGTGCTGCTGGAGGTGACAGCTTTA 450
  |||||||
134 alaIleuAlaGlnLeuValValMetValleuValGlnValThrAlaLeu 150
451 GGCACCTGAGGATGCTCATCAGTAATATCTTCAACAGACTACACAT 500
  |||||||
151 GlyThrleuArgMetValIleSerAsnIlePheasnIrrpIrrhSme 167
501 GAAATGATGCACATCTCACTGTTCGAGCCCTATTTTGGGCTGTGTGG 550
  |||||||
167 TAsnleuArgHisPheTyrValPheAlaIaIaTyrPheOlyleuThrValA 184
551 CCTGTGCTGCCCAAGCTTCAACTCTGCTGTGAGAGATCCATCGAA 600
  |||||||
184 latrpCysleuProIlySerProleuProlYsGlyThnGlnaspAspGln 200
601 ACAGCAGATRACCAAGTTGTGTGCTGCTGCGGCGCCCTTCTGTG 650
  |||||||
201 ArgAlaThrIleProSerleuSerAlaMetleuGlyAlaIaIaIaLeu 217
651 GATGTTTGGCCCAAGCTTCAACTCTGCTGTGAGAGATCCATCGAA 700
  |||||||
217 PheThrIrrpProSerValAsnSerProleuIleuArgSerProIleGlnA 234
701 GGAGATATGCCGTGTTCACACCTACTATGTGTGACAGTACGCTGTG 750
  |||||||
234 rGlyAsnAlaMetPheAsnThrTyrTyrAlaIaIaIaValSerValAl 250
751 ACAGCATCTCAGGTCATCTTGTGCTGCTGACCCCAAGGAGATCAGCA 800
  |||||||
251 ThrAlaIleSerGlySerSerleuAlaHisProGlnArgLysIleSerne 267
801 GACTTATGTGCACAGTCCGCTGTGGCAGAGCGCTGTGGTGTACCT 850
  |||||||
267 tThrTyrValHisSerAlaValleuAlaGlyValAlaValAlaGlyThrs 284
851 CGTGTACCTGATCCCTTCCGTGCTGCTGCTGCTGCTGCTGCTGCTG 900
  |||||||
284 eCysHisIleuIleProSerProIrrpLeuAlaMetValleuGlyleuVal 300
901 GCTGGGCTGATCTCCGTCGCGGGGAGCCAACTACCTCCGGGCTGTGTA 950
  |||||||
301 AlaGlyleuIleSerIleGlyGlyAlaLysCysLeuProValCysCysAs 317
951 CCGAGTCTGGGATTCGCCACAGCTCCATCATGAGCTACAACTCAGCT 1000
  |||||||
317 PArgValleuGlyIleHisHisIleSerValMetHisSerIlePheSerL 334
1001 TGTGTGCTGTCTGAGAGATCATCTACATGTGTGCTGCTGCTGTGAT 1050
  |||||||
334 euIleuGlyleuGlyGlnIleThrTyrIleValleuIleuValleuHis 350
1051 ACCGTGCGAGCCGCAATGGCATGATGGCTCCAGGCTCCCTCAGCAT 1100
  |||||||
351 ThrValIrrpAsnGlyAsnGlyMetIleGlyPheIleValIleuIleuSerI 367
1101 TGGGAACTCAGTGTGCGCATCTGATAGCTTCTCAGCTGTGCTGCTGA 1150
  |||||||
367 eGlyGlnleuSerleuAlaIleValIleAlaIeThrSerGlyleuLeuT 384
1151 CAGGTTGCTCTAAATCTTAATATGAGAAAGACCTCATGAGGCTAA 1200
  |||||||
384 hrGlyleuIleuIleuAsnleuLysIleTrrpYsAlaIrrpHisValAlaLys 400
1201 TATTTTGTATGACCAAGTTTCTGAGAGTTCCTCATTTGGCTGTGAGTT 1250
  |||||||
401 TyrPheAspArgGlnValPheTrrpYsPheTrrpHisIleuAlaValGlyPh 417
1251 T 1251
417 e 417

```

THIS PAGE BLANK (USPTO)

A: Molecule type: mRNA
 A: Residues: 1417 <RES>
 A: Cross-references: GB:U37052; NID:9607009; PIDN:AAA65626.1; PID:9607010

alignment_scores:
 Quality: 1908.00 Length: 417
 Ratio: 4.806 Gaps: 0
 Percent Similarity: 95.204 Percent Identity: 89.928

alignment_block:
 US-09-600-714-41 x I37075

Align seg 1/1 to: I37075 from: 1 to: 417

```

1 ATGACCTTAAGTACCCGCGGTCTGTCCGCGCTGCGCTGCGCTGCGG 50
|||||
1 MetSerSerLysTyrProArgSerValArgCysCysLeuProLeuLysAl 17
51 CCTAACACTGGAAAGCTCTCATCTCTCTCTCTCTCTCTCTCTCTCT 100
|||||
17 AleuThrLeuGluAlaAlaLeuLeuLeuLeuPheTyrPhePheThr 34
|||||
101 ATGACCTTCCTTAGAGATCAAAAGGCGCTGCGCATCTCTCAAGTT 150
|||||
34 YAspAlaSerLeuGluAspGlnGlnGlyLeuValAlaSerTyrGlnVal 50
|||||
151 GGCGAAGATCTGACCGTGAATGGCGCCATTGGCTGGGGTCTCTAC 200
|||||
51 GlyGlnAspLeuThrValMetAlaAlaIleGlyPheGlyPheLeuThr 67
|||||
201 GAGTTCCGGAGACACAGCTGGAGCGATGGCGCTCAACCTCTTCATG 250
|||||
67 rSerPheAlaArgHisSerTyrPheSerValAlaPheAsnLeuPheMet 84
|||||
251 TGCGCGCTGGGTGTCAGTGGGCAATCCTGCTGACGCGCTTCTGAG 300
|||||
84 euAlaLeuGlyValGlnTyrAlaIleLeuLeuAspGlyPheLeuSerGln 100
|||||
301 TTCCTCTTGGGAAGGTGTCATCACATGTCTCAATATTCGCTGGCC 350
|||||
101 PheProProGlyLysValValIleTyrLeuPheSerIleArgLeuAla 117
|||||
351 CATGAGTCTTGTGCGTGTGATCTCAATGATGCTGCTCTGGGGAAG 400
|||||
117 rMetSerAlaLeuSerValLeuLeuSerAlaGlyAlaValLeuGlyTyr 134
|||||
401 TCAACTGGGCGAGTGTGCTGATGGTGGTGGGAGGAGTACAGCTTT 450
|||||
134 AlAsnLeuValGlnLeuValValMetValLeuValGlnValThrAlaLeu 150
|||||
451 GGCAACCTGGAGATGACATCAGTAAATATCTTCAACACAGACTACAT 500
|||||
151 GlyThrMetArgMetValIleSerAsnIlePheAsnThrAspTyrHis 167
|||||
501 GAACATATGCAATCTACGCTGTTCGACGCTATTTGGCTGTCTGG 550
|||||
167 tAsnMetThrHisPheTyrValAlaPheAlaAlaTyrPheGlyValThrVal 184
|||||
551 CCTGCTGCTGGCCAAAGCCTTACCCGCGGAGGAGGAGGATTAAGAT 600
|||||
184 lArpProGlyLeuProGlyProLeuProAspIleLysGlnAspLysAspGln 200
|||||
601 ACACCAAGATACCCAGTTTGTCTGCATGCTGGGCGCCCTCTCTTGT 650
|||||
201 lLeuAlaThrIleProSerLeuSerAlaMetLeuGlyThrLeuPheLeu 217
|||||
651 GATGTTGGCCAAAGTTCAACTGCTCTGTCTGGAAGTCAATCGAA 700
|||||
217 pMetPheThrProSerPheAsnSerAlaLeuLeuArgSerProIleGlu 234
|||||
701 GGAGAAAGCCGCTTCAACACTTACTGCTGACGATGAGCTGGTG 750
|||||

```

```

234 rGlyAsnAlaValAlaPheAsnThrTyrTyrAlaLeuAlaValSerValVal 250
|||||
751 ACAGCCATCTCAGGCTCATCTTGGCTCACCCCAAGGAAATCAGCA 300
|||||
251 ThrAlaIleSerValSerSerLeuAlaHisProGlnGlyLysIleAsn 267
|||||
801 GACTATGTGCACAGTGGCGTGTGGCAGAGGCGCTGCTGGTGA 350
|||||
267 tThrTyrMetHisAsnAlaValLeuAlaGlyGlyValAlaValGlyThr 284
|||||
851 CGTGCACCTGATCCCTCTCCGCTGGCGATGGCATGGCTGGCTGTG 900
|||||
284 eTyrHisIleLeuIleThrSerProThrPheAlaMetValLeuGlyLeuVal 300
|||||
901 GCTGGGCTGATCTCCGTCGCGGAGCCAAAGTACCTGCGGGGTGTGTA 950
|||||
301 AlaGlyLeuIleSerIleGlyGlyAlaLysGlyLeuProGlyCysCys 317
|||||
951 CCGAGTGTGGGGAATCCCAAGCTCATCTGAGGCTTCAACTGAGT 1000
|||||
317 nArgValLeuGlyIleHisAspSerSerValMetHisTyrAsnPheSer 334
|||||
1001 TGCTGGCTGCTGGAGAGATCATCTACATGTGCTGCTGCTGAT 1050
|||||
334 euLeuGlyLeuLeuGlyGlyIleThrTyrIleValLeuMetValLeuHis 350
|||||
1051 ACCGTGCGAGCCGCAATGCAATGATGGCTTCCAGCTCTCTCAAGAT 1100
|||||
351 ThrValGlyAlaGlyAsnGlyMetValGlyPheGlnValLeuValSer 367
|||||
1101 TGGGGAAGTCACTTGGCATGCTGATACCTGCAAGCTGGCTGCTG 1150
|||||
367 rGlyGlyLeuSerLeuAlaLeuAlaIleValAlaThrSerGlyLeuMet 384
|||||
1151 CAGGTTCTCCCTAAATCTTAAATGGAAGCACTGATGAGGCTTAA 1200
|||||
384 hTrpIleuLeuLeuAsnLeuLysIleTyrLysAlaProHisAlaLys 400
|||||
1201 TAAATTGATGACCAAGTTTCTGAAATTTCTCATTTGGCTGTGAT 1250
|||||
401 tYrPheAspAspGlnValPheTyrLysPheProHisLeuAlaValGly 417
|||||
1251 T 1251
417 e 417

```

seq_name: pI2:184434

seq_documentation_block:
 Rhesus-like protein - crab-eating macaque
 C:Species: Macaca fascicularis (crab-eating macaque)
 C:date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 21-Jul-2000
 C:Accession: I84434
 R:Salvignol, I.; Blancher, A.; Calvus, P.; Clayton, J.; Socha, W.W.; Collin, Y.; Ruffi
 Blochem. Genet. 32, 201-221, 1994
 A:Title: Molecular genetics of chimpanzee Rh-related genes: their relationship with t
 A:Reference number: I37003; MUID:9508595
 A:Accession: I84434
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-417 <RES>
 A:Cross-references: GB:U37054; NID:9609515; PIDN:AAA65628.1; PID:9609516

alignment_scores:
 Quality: 1639.00 Length: 417
 Ratio: 4.313 Gaps: 0
 Percent Similarity: 91.127 Percent Identity: 78.177

alignment_block:

US-09-600-714-41 x I84434
 Align seg 1/1 to: I84434 from: 1 to: 417

THIS PAGE IS ANK (USPTO)


```

1 ATGAGCTCTAAGTACCCGCGTCTGTCCGGCGCTGCCCTCCCTGGGC 50
1 MetSerSerLysTyrProArgSerValArgCysCysLeuProLeuTrpAl 17
51 CCTACACTGAGACAGCTCTCATCTCTCTTCTATTTTACCACACT 100
17 aLeuThrLeuGlnAlaLeuIleLeuLeuPhePhePheThrTyrT 34
101 ATGAGCGTCTCTAGAGATCAAAAGGGCGTGGCGCTCTCATCAAGTT 150
34 YrAPhAlaSerLeuGlnAspGlnValValAlaSerTyrGlnVal 50
151 GGCCAGATCTGACCGGTGATGGCGCCATTGGCTGGCTCTCCACCTC 200
51 CysGlnAspLeuThrValMetAlaValLeuGlyLeuGlyPhePheThrSe 67
201 GAGTTTCCGGAGACACAGCTGGACAGTGTGGCTTCAACTCTTCATGC 250
67 rAsnLeuArgArgAsnSerTyrSerValAlaPheAsnLeuPheLeuL 84
251 TGGCGCTTGTGTGACAGTGGCAATCCGTGGACGGCTTCTGAGCCAG 300
84 euAlaLeuGlyValGlnTrpAlaIleLeuLeuAspGlyPheLeuSerGln 100
301 TTCCCTCTGGGAAGTGTGCATCACACTGTTCAGTATTCGGCTGGCCAC 350
101 PheSerProGlyLysValIleValIleLysLeuPheSerIleArgLeuAlaThr 117
351 CATGAGTCTTTCGCGTGTGATCTCAGTGGAGTGTCTCTGGGGAAG 400
117 rArgSerThrThrSerMetLeuIleSerMetAsnAlaValLeuGlyLysV 134
401 TCAACTTGGCCAGTGTGTGTGATGTGTGTGGAGAGTGTGACAGCTTA 450
134 alAlaLeuAlaGlnLeuValValMetGlnLeuValGlnLeuValPhe 150
451 GGCACTGAGAGATGTGCATCATATATCTTCAACACAGACTACACACT 500
151 GlyThrMetArgIleValIleTyrAsnIlePheLysIleAspTyrGlyMe 167
501 GAACATGATGCACATCTACGTCGTGGCGCACTATTTGGCTGCTGGG 550
167 tAsnMetGlnHisIleHisValPheAlaAlaTyrPheGlyLeuThrValA 184
551 CCTGTGCTCCCAAGGCTCTACCCGAGGAGGAGAGATAAAGATCAG 600
184 lAlrPcysLeuProLysProLeuProLysGlyThrGlnAspLysTyrGln 200
601 ACAGCAACGATACCCAGTTTGTGTCGCATGCTGGCGCCCTCTCTGTG 650
201 ThrThrThrSerProSerLeuPheAlaMetLeuGlyThrLeuPheLeuTr 217
651 GATGTTTGGGCAAGTTCACACTCTGCTGCTGCTGAGAAAGTTCATCGAA 700
217 pMeCherThrProThrPheAsnSerAlaLeuLeuAsnProIleGlnA 234
701 GGAAGATGCGGTGTCAACACTACTATGCTAGCAGTACAGCGGTG 750
234 rGlyAsnAlaValPheSerThrTyrTrpAlaLeuAlaValSerAlaVal 250
751 ACAGCATCTCAGGATCATCTTGGCTACCCCAAGGAGATCAGCAA 800
251 ThrAlaIleSerValSerSerLeuAlaHisProGlnArgLysIleAsnMe 267
801 GACTTATGTGACAGTGTGGTGTGGCGAGGAGGCTGCTGGTGTACT 850
267 tThrTyrMetProAsnAlaGlyLeuAlaGlyValAlaValGlyAlaAs 284
851 CGGTGACCTATCCTTCCGCGTGTGGAGTGTGGAGTGTGGTGTGTG 900
284 erGysHisValIleHisSerProTrpIleAlaMetValLeuGlyLeuVal 300

```

```

901 GCTGGGCTGATCTCCGCGGGGAGCCCAAGTACCTGCCGGGGTGTGTAA 950
301 AlAlaIleLeuIleSerPheGlyAlaLysCysLeuProValCysPheAs 317
951 CCGAGTCTGGGATTCGCCACAGCTTCATCATGCGGTACACTTACGT 1000
317 nArgValLeuGlyIleHisGlnSerHisSerMetHisTyrThrPheGlyL 334
1001 TGCTGGCTGCTGCTGAGAGATCATCATATGCTGCTGCTGCTGAT 1050
334 euProAlaLeuLeuGlyGlnIleThrTyrIleValAlaMetAlaLeuArg 350
1051 ACCGTGGAGCCGGCAATGCGCATGATGCTTCCAGGCTCTCTCAGCAT 1100
351 ValPheThrPheAlaSerSerAsnMetIleGlyPheGlnValLeuLeuSerTh 367
1101 TGGGAGCATCAGCTTGGCGCATCGATAGCTCCAGCTGCTGCTCTGA 1150
367 rGlyThrLeuSerLeuAlaMetAlaMetSerIleThrSerGlyLeuLeuT 384
1151 CAGGTTGCTCTCTTAATCTTAAATATGAAAGACCTCATGAGGCTTAA 1200
384 hrGlyLeuLeuLeuAsnLeuLysIleTrpLysGlyProHisValAlaLys 400
1201 TATTTGATGACCAAGTTTCTGGAAGTTTCTCATTTGCTGTGGATT 1250
401 TyrPheAspAspGlnAlaPheTrpGlnPheProHisLeuAlaValGlyPh 417
1251 T 1251
417 e 417

```

seq_name: p1r2:S78480

seq_documentation_block:

Rhesus blood group antigen-like protein isoform 4 - human
 C:Species: Homo sapiens (man)
 C:date: 11-Mar-1998 #sequence_revision 17-Apr-1998 #text_change 05-Nov-1999
 C:accession: S78480; S26560
 R:Colin, Y.
 submitted to the EMBL Data Library, November 1991
 A:Reference number: S78478
 A:Accession: S78480
 A:Molecule type: mRNA
 A:Residues: 1-354 <COL>
 A:Cross-References: EMBL:X63098; NID:936019; PID:CAA4812.1; PID:936020
 R:Le Van Kim, C.; Cherif-Fahar, B.; Reynal, V.; Mouuo, I.; Lopez, M.; Carton, J.F.;
 Blood 80, 1074-1078, 1992
 A:Title: Multiple Rn messenger RNA isoforms are produced by alternative splicing.
 A:Reference number: S26560; MUID:92360855
 A:Accession: S26560
 A:Molecule type: mRNA
 A:Residues: 160-354 <LEV>
 A:Cross-References: EMBL:X63098
 A:Experimental source: tissue-type bone marrow
 C:Keywords: alternative splicing; erythrocyte; surface antigen; transmembrane protein

alignment_scores:

Quality:	Ratio:	Length:
1627.00	4.771	399
Percent Similarity:	85.464	Percent Identity:
		82.206

alignment_block:

US-09-600-714-41 x S78480

Align seg 1/1 to: S78480 from: 1 to: 354

```

1 ATGAGCTCTAAGTACCCGCGTCTGTCCGGCGCTGCCCTCCCTGGGC 50
1 MetSerSerLysTyrProArgSerValArgCysCysLeuProLeuTrpAl 17
51 CCTACACTGAGACAGCTCTCATCTCTCTTCTATTTTACCACACT 100

```

THIS PAGE BLANK (USPTO)


```

|||||
Db 91 GAAGAGCTCATCTCCCTCTCTATTTTATTTTACCACATAGACCCTCTTAGAGAT 150
121 caaaaggcctctgacatccataccaattggcgaacatctgacgctgtgtggccatt 180
151 caaaaggcctctgacatccataccaattggcgaacatctgacgctgtgtggccatt 210
181 ggccttgagctccctcaccctcgaatttcaggagacacagcttgagcagcttgagcctc 240
211 ggccttgagctccctcaccctcgaatttcaggagacacagcttgagcagcttgagcctc 270
241 cctctcattcctgagcctctgctgctgctgagcttgagcagcttgagcagcttgagc 300
271 cctctcattcctgagcctctgctgctgctgagcttgagcagcttgagcagcttgagc 330
301 ttcctctcggagaggttgctcaccacacgcttgacgcttgagcagcttgagcagcttg 360
331 ttcctctcggagaggttgctcaccacacgcttgacgcttgagcagcttgagcagcttg 390
361 tctgctgagctgacatccatcagctgagctgctgctgagagagctcacttgagcagct 420
391 atgtcgtgctgacatccatcagctgagctgctgctgagagagctcacttgagcagct 450
421 gtagatgctgctgagctgagctgagcttgagcagcttgagcagcttgagcagcttg 480
451 gtagatgctgctgagctgagctgagcttgagcagcttgagcagcttgagcagcttg 510
481 ttcacacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 540
511 ttcacacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 570
541 ctgctctgagcctgagctgagcagcagcagcagcagcagcagcagcagcagcagcag 600
571 ctgactgtgctgagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 630
601 acagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 660
631 agagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 690
661 ccaagcttcaactctgctgctgagagagctcagcagcagcagcagcagcagcagcag 720
691 ccaagcttcaactctgctgctgagagagctcagcagcagcagcagcagcagcagcag 750
721 acctactatgctgtagcagcagcagcagcagcagcagcagcagcagcagcagcagc 780
751 acctactatgctgtagcagcagcagcagcagcagcagcagcagcagcagcagcagc 810
781 ccccaaggagagagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 840
811 ccccaaggagagagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 870
841 gtaggagcctgctgtagcagcagcagcagcagcagcagcagcagcagcagcagcagc 900
871 gtaggagcctgctgtagcagcagcagcagcagcagcagcagcagcagcagcagcagc 930
901 gctgagcctgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 960
931 gctgagcctgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 990
961 ggcgtgagcctgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1020
991 ggcgtgagcctgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1050
1021 atcaatcatatgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 1080
1051 atcaatcatatgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 1110
1081 ttcacagcctcctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 1140
1111 ttcacagcctcctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 1170
1141 gctcctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1200
|||||

```

```

Db 1171 GGTCTCCTGACAGTTTGTCTCTAATCTCAAAATATGGAAGACACCTCATGTGCTAA 1230
1201 tatcttgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 1254
1231 tatcttgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 1284

RESULT 2
US-08-553-888A-2
; Sequence 2, Application US/08553888A
; Patent No. 5723293
; GENERAL INFORMATION:
; APPLICANT: Huang
; TITLE OF INVENTION: DIAGNOSTIC METHOD AND KIT FOR
; TITLE OF INVENTION: DETERMINING RH BLOOD GROUP GENOTYPE
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hoffmann & Baron, LLP
; STREET: 350 Jericho Turnpike
; CITY: Jericho
; STATE: New York
; COUNTRY: USA
; ZIP: 11753
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 MB storage
; COMPUTER: IBM compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/553,888A
; FILING DATE: 11/06/95
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: O'Dea, Sean W.
; REGISTRATION NUMBER: 37690
; REFERENCE/DOCKET NUMBER: 454-5
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 822-3550
; TELEFAX: (516) 822-3582
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1466 nucleotides
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; US-08-553-888A-2

Query Match 94.3%; Score 1182; DB 1; Length 1466;
Best Local Similarity 96.4%; Pred. No. 0;
Matches 1209; Conservative 0; Mismatches 45; Indels 0; Gaps 0;

1 atagagctcaagtaacccggagctgctcggagcagcagcagcagcagcagcagcagcagc 60
45 atagagctcaagtaacccggagctgctcggagcagcagcagcagcagcagcagcagcagc 104
61 gaagagcctcaatcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctc 120
105 GAAGAGCTCTATCTCTCTCTATTTTATCCCATATGAGCGCTCTCTTAGAGAT 164
121 caaaaggcctgtagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 180
165 caaaaggcctgtagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 224
181 ggccttgagcctcctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 240
225 ggccttgagcctcctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 284
241 cctctcattcctgagcctctgctgctgctgagcagcagcagcagcagcagcagcagcag 300
285 cctctcattcctgagcctctgctgctgctgagcagcagcagcagcagcagcagcagcag 344
301 ttcctctcggagaggttgctcaccacacgcttgacgcttgagcagcagcagcagcagcag 360

```

```

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 345 TTCCCTCCTGGAGGAGGTGATCATCACTGTTAGTATTCGGTGGCCACCATATGAGTCT 404
Qy 361 ttctcgctgctgatactcaatgagatgctgctctgggaaagtcacactggcgagctgtg 420
Db 405 ATGTCGGTGGTGAATCAGCGGGGCTGCTCTTGGGAAAGGCAACCTGGCCACTTGGTG 464
Qy 421 gtgattgtgctgctgagaggtgacagcttaagcaacctgaagatgatacagtaataac 480
Db 465 GTGATGCTGGTGGTGGAGGTACAGCTTTAGGCACTTGAAGAGTGCATATGTAATATC 524
Qy 481 ttcaacacagataccacatagaataatgacacatacagtgctgcagacctatcttg 540
Db 525 TTCAACACAGATACCAATGAACTGAGGCACTTACGTACGTGCGAGCCTATTTGGG 584
Qy 541 ctgctgtgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 600
Db 585 CTGACTGTGGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 644
Qy 601 acaagcaagataccagcttctgctgcaatgctgggcccctctctgtgattctg 660
Db 645 AGACCAAGATACCACTGTTTCTGTCATGCTGGGCGCCCTCTTGTGTGATGTTCTGG 704
Qy 661 ccaagttcaactctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 720
Db 705 CCAAGTGTCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 764
Qy 721 accaactatgctgatacagatgctgctgctgctgctgctgctgctgctgctgctgct 780
Db 765 ACCTACATGCTCTAGAGTACAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 824
Qy 781 ccccaaggaagatcaagaagactatgctgctgctgctgctgctgctgctgctgctgctg 840
Db 825 CCCCAGGAAGAGATCAGCATGACTTATGTGACAGTGGGCTGTGGCAGGAGGCTGCT 884
Qy 841 gtgggtacactcgtgtcaactgataccctctcctgctgctgctgctgctgctgctgctg 900
Db 885 GTGGGTACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 944
Qy 901 gctggtgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 960
Db 945 GCTGGGTGATCTCATCTGCGGGGAGCCAAAGTCCCTGCGGTGTGTGTGTGTGTGTGTGT 1004
Qy 961 gggattccccagagctcatatgctgctgctgctgctgctgctgctgctgctgctgctg 1020
Db 1005 GGGATTACCACTCTCCGTATGACTCCATCTTCACTGCTGGGCTCTCTCTGAGAG 1064
Qy 1021 atcaactaactgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 1080
Db 1065 ATCACTGATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1124
Qy 1081 ttcaagctctctcctcagcatggtggaactcaagcttgccatcgtgatagctcactgct 1140
Db 1125 TTCCAGGCTCCCTCAGATGGGGAACCTGACTGTGGCCATGCTGATGCTCTCAGCTCT 1184
Qy 1141 ggtctcctgaaggtgtgctcctaaatctaaatgaataatgaagacacctcagaagctaa 1200
Db 1185 GGTCTCCGACAGGTTGCTCTTAATCTCAAAATATGGAAGCACTCATGTGGCTTAA 1244
Qy 1201 tattttgatgacaaagtttctgaaagtttccatctgctgctgctgctgctgctgct 1264
Db 1245 TATTTGATGACCAAGTTTCTGGAAGTTTCCATTTGGCTGTGTGATTTTAA 1288

```

```

RESULT 3
US-08-476-176B-15/c
: Sequence 15, Application US/08476176B
: Patent No. 5958708
: GENERAL INFORMATION:
: APPLICANT: Hardman, No. 5958708man
: APPLICANT: Koldinger, Frank
: APPLICANT: Saldanha, Jose
: TITLE OF INVENTION: Reshaped monoclonal antibodies against an

```

```

: TITLE OF INVENTION: Immunoglobulin isotype
: NUMBER OF SEQUENCES: 55
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: No. 5958708artis Patent Department
: STREET: 59 Route 10
: CITY: East Hanover
: STATE: New Jersey
: COUNTRY: USA
: ZIP: 07936-1080
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/476,176B
: FILING DATE:
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/127,721
: FILING DATE: 27-SEPTEMBER-1993
: APPLICATION NUMBER: US 07/952,802
: FILING DATE: 25-SEPTEMBER-1992
: ATTORNEY/AGENT INFORMATION:
: NAME: No. 5958708ak, Henry P.
: REGISTRATION NUMBER: 33,200
: REFERENCE/DOCKET NUMBER: 4-19276/A/P2/CIP
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (908) 277-5110
: TELEFAX: (908) 277-4306
: INFORMATION FOR SEQ ID NO: 15:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 467 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
: MOLECULE TYPE: DNA (genomic)
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 22..447
: FEATURE:
: NAME/KEY: mat_peptide
: LOCATION: 79..447
: OTHER INFORMATION: /product= "heavy chain variable
: OTHER INFORMATION: region C21-Hay1"
US-08-476-176B-15

Query Match 3.1%; Score 39.4; DB 2; Length 467;
Best Local Similarity 51.4%; Pred. No. 0.019;
Matches 91; Conservative 0; Mismatches 86; Indels 0; Gaps 0;

Qy 264 gcaatgggcaactcctgctgagcgtcttcctgagccagttccctctggaagtgatcat 323
Db 405 GTACTGCTAGTTGCTGCGCGTGAAGTGCTGAACCTGGCGCAGTAGTAACGCGGGTGC 346
Qy 324 caactgttcgtatcttgctgctgctgctgctgctgctgctgctgctgctgctgctgct 383
Db 345 CTCGCTGCTCAGGCTGCTCAGCTCCATGTAAGCGCTGCTGCTGCTGCTGCTGCTGCTG 386
Qy 384 tgcgtctggggaagtcacacttggcgcaatggtggtggtggtggtggtggtggtggtg 440
Db 285 GGTGGCTTGGCTTGAACCTTCTGTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 229

```

```

RESULT 4
US-08-127-721A-15/c
: Sequence 15, Application US/08127721A
: Patent No. 6066718
: GENERAL INFORMATION:
: APPLICANT: Hardman, No. 6066718man
: APPLICANT: Koldinger, Frank
: APPLICANT: Saldanha, Jose

```

```

1 TITLE OF INVENTION: Reshaped monoclonal antibodies against an
2 TITLE OF INVENTION: Immunoglobulin isotype
3 NUMBER OF SEQUENCES: 55
4 CORRESPONDENCE ADDRESS:
5 ADDRESSSEE: No. 606671Bartis Patent and Trademark Department
6 STREET: 59 Route 10
7 CITY: East Hanover
8 STATE: New Jersey
9 COUNTRY: USA
10 ZIP: 07936-1080
11 COMPUTER READABLE FORM:
12 MEDIUM TYPE: Floppy disk
13 COMPUTER: IBM PC compatible
14 OPERATING SYSTEM: PC-DOS/MS-DOS
15 SOFTWARE: Patentin Release #1.0, Version #1.25
16 CURRENT APPLICATION DATA:
17 APPLICATION NUMBER: US/08/127,721A
18 FILING DATE: 27-SEPTEMBER-1993
19 CLASSIFICATION: 424
20 PRIOR APPLICATION DATA:
21 APPLICATION NUMBER: US 07/952,802
22 FILING DATE: 25-SEPTEMBER-1992
23 ATTORNEY/AGENT INFORMATION:
24 NAME: No. 6066718ak, Henry P.
25 REGISTRATION NUMBER: 33,200
26 REFERENCE/DOCKET NUMBER: 4-19276/A/P2/CIP
27 TELECOMMUNICATION INFORMATION:
28 TELEPHONE: (908) 277-5110
29 TELEFAX: (908) 277-4306
30 INFORMATION FOR SEQ ID NO: 15:
31 SEQUENCE CHARACTERISTICS:
32 LENGTH: 467 base pairs
33 TYPE: nucleic acid
34 STRANDEDNESS: double
35 TOPOLOGY: linear
36 MOLECULE TYPE: DNA (genomic)
37 FEATURE:
38 NAME/KEY: CDS
39 LOCATION: 22..447
40 FEATURE:
41 NAME/KEY: mat_peptide
42 LOCATION: 79..447
43 OTHER INFORMATION: /product= "heavy chain variable
44 region C21-hay1"
45 US-08-127-721A-15
46
47 Query Match 3.1%; Score 39.4; DB 3; Length 467;
48 Best Local Similarity 51.4%; Pred. No. 0.019;
49 Matches 91; Conservative 0; Mismatches 86; Indels 0; Gaps 0
50
51 Oy 264 gcaatggagcaaatcctgcttgacgcgttcctgaagcagttcccttctggaaagtgtgat 323
52 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
53 Db 405 gtatgcgtatgattgtcgccgctgaactgacctgaacctgaccgcatgatacgcggctgtc 346
54 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
55 Oy 324 cacactgttcacgatcttgagctgacacatgtagtcttgttcgtagtcatcatcagttga 383
56 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
57 Db 345 ctccgttgctcacgctgctcctcactatgaggcggctgctgccccttgcttccggcgtgaa 286
58 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
59 Oy 384 tgcctcttgaggaaagtcacttgccagcttggtgtagtgcgtgtgctgtgtaagt 440
60 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
61 Db 285 ggtgcgcttgcccttgaaacttctcgcttgagtggtggcgtgaagtcgccggcgctgat 229

```

```

TITLE OF INVENTION: immunoglobulin isolate
NUMBER OF SEQUENCES: 55
CORRESPONDENCE ADDRESS:
ADDRESSSEE: No. 6072035artis Patent Department
STREET: 59 Route 10
CITY: East Hanover
STATE: New Jersey
COUNTRY: USA
ZIP: 07936-1080

COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/485,246A
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/127,721
FILING DATE: 27-SEPTEMBER-1993
APPLICATION NUMBER: US 07/952,802
FILING DATE: 25-SEPTEMBER-1992
ATTORNEY/AGENT INFORMATION:
NAME: No. 6072035ak, Henry P.
REGISTRATION NUMBER: 33,200
REFERENCE/DOCKET NUMBER: 4-19276/P2/CIP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (908) 277-5110
TELEFAX: (908) 277-4306
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 467 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 22..447
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 79..447
OTHER INFORMATION: /product= "heavy chain variable"
OTHER INFORMATION: region C21-Hay1"
US-08-485-246A-15

Query Match          3.1%; Score 39.4; DB 3; Length 467;
Best Local Similarity 51.4%; Pred. No. 0.019; 86; Indels 0; Gaps 0
Matches 91; Conservative 0; Mismatches 66; Indels 0; Gaps 0

QY      264   gcaagtggagcaactcgtctgacgcgttccttcaggccagttcccttctggaaagtgatcat 323
           |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db       405   gTAGTCGTAGTTGCTGCACGGCTGAAGTGCGTGAACCTGGCGCAGTAGTACACGGCGGTGTC 346

QY      324   cacactgttcagtatctggcgctgaccacacatgatgcttgttcgtygctgatctcagtgga 383
           |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db       345   CTCGGTGTGCTGACGGCTGCTCACGCTCATGATGAGCGGTGCTGGCGGTGCTGCGGGGTGAA 286

QY      384   tgcctctctggaaaggtccaacttgagcgagctgtgtgtgtagtgcgtgtgagaagt 440
           |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db       285   GGTCGCCCTTGCGCTTGACTTCTCGTTGTAGTTGGTGTGAAGTGCCCGGGGCTGAT 229

RESULT
6
US-08-476-176B-11/c
Sequence 11, Application US/08476176B
Patent No. 5958708
GENERAL INFORMATION:
APPLICANT: Hardman, No. 5958708man
APPLICANT: Kolbinger, Frank
APPLICANT: Saldanha, Jose
```

```

1  TITLE OF INVENTION:      Reshaped monoclonal antibodies against an
2  TITLE OF INVENTION:      immunoglobulin isotype
3  NUMBER OF SEQUENCES:    55
4  CORRESPONDENCE ADDRESS:
5  ADDRESSEE: No. 5958708artis Patent Department
6  STREET: 59 Route 10
7  CITY: East Hanover
8  STATE: New Jersey
9  COUNTRY: USA
10 ZIP: 07936-1080
11
12 COMPUTER READABLE FORM:
13 MEDIUM TYPE: Floppy disk
14 COMPUTER: IBM PC compatible
15 OPERATING SYSTEM: PC-DOS/MS-DOS
16 SOFTWARE: PatentIn Release #1.0, Version #1.25
17 CURRENT APPLICATION DATA:
18 APPLICATION NUMBER: US/08/476,176B
19 FILING DATE:
20 CLASSIFICATION: 435
21 PRIOR APPLICATION DATA:
22 APPLICATION NUMBER: 08/127,721
23 FILING DATE: 27-SEPTEMBER-1993
24 APPLICATION NUMBER: US 07/952,802
25 FILING DATE: 25-SEPTEMBER-1992
26 ATTORNEY/AGENT INFORMATION:
27 NAME: No. 5958708ak, Henry P.
28 REGISTRATION NUMBER: 33,200
29 REFERENCE/DOCKET NUMBER: 4-19276/R/P2/CIP
30 TELECOMMUNICATION INFORMATION:
31 TELEPHONE: (908) 277-5110
32 TELEFAX: (908) 277-4306
33 INFORMATION FOR SEQ ID NO: 11:
34 SEQUENCE CHARACTERISTICS:
35 LENGTH: 468 base pairs
36 TYPE: nucleic acid
37 STRANDEDNESS: double
38 TOPOLOGY: linear
39 MOLECULE TYPE: DNA (genomic)
40 FEATURE:
41 NAME/KEY: CDS
42 LOCATION: 22..447
43 FEATURE:
44 NAME/KEY: mat.peptide
45 LOCATION: 79..447
46 OTHER INFORMATION: /product= "heavy chain variable
47 OTHER INFORMATION: region C21-H1"
48
49 US-08-476-176B-11

```

Query Match	3.1%:	Score 39.4:	DB 2:	Length 468:
Best Local Similarity	51.4%:	Pred. No. 0.019:	Mismatches 86:	Indels 0:
Matches 91:	Conservative	0:	Mismatches 86:	Indels 0:
QY 264	gcacgtggcaacatccgtctgcagagcgtccctcgaacagcttccctctcgtggaagtgtgcat	323		
Db 405	GTAGCGGTAAGTTGGCTGGCCGGCTGAAGTGGCTGAACCTGGCGCGCAGTAGTACACGGCGGTGC	346		
QY 324	caaacctgtgcatttcggctcgtgcacacacatgagttccttttcgtcgtatctcagatga	383		
Db 345	CTCCCTGGTCAAGCGTGCACACTCCATATGAAGCGGTGTGGTGGTGGTGGCGCGTGA	286		
QY 384	tgcctgtcttgggaagatgcacacttggcgcagttgtgtgtatgctgcgtggaagt	440		
Db 285	GGTGGCCCTTGCCCTTGAACCTTCTCGTTATAGTTGGTGGTGAAGGTCGGCGGCGCTAT	229		

RESULT: 7
US-08-127-721A-11/C
: Sequence 11, Application US/08127721A
: Patent No. 6066718
: GENERAL INFORMATION:
: APPLICANT: Hardman, No. 6066718man
: APPLICANT: Kolbinger, Frank

APPLICANT: Saldanha, Jose
 TITLE OF INVENTION: Reshaped monoclonal antibodies against an
 TITLE OF INVENTION: Immunoglobulin Isotype
 NUMBER OF SEQUENCES: 55
 CORRESPONDENCE ADDRESS:
 ADDRESS: No. 6066718artis Patent and Trademark Department
 STREET: 59 Route 10
 CITY: East Hanover
 STATE: New Jersey
 COUNTRY: USA
 ZIP: 07936-1080
 COMPUTER READABLE FORM:
 MEDIUM TYPE: floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/127,721A
 FILING DATE: 27-SEPTEMBER-1993
 CLASSIFICATION: 424
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/952,802
 FILING DATE: 25-SEPTEMBER-1992
 ATTORNEY/AGENT INFORMATION:
 NAME: No. 6066718ak, Henry P.
 REGISTRATION NUMBER: 33,200
 REFERENCE/DOCKET NUMBER: 4-19276/A/P2/CIP
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (908) 277-5110
 TELEFAX: (908) 277-4306
 INFORMATION FOR SEQ ID NO: 11:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 468 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: DNA (genomic)
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 22..447
 FEATURE:
 NAME/KEY: mat_peptide
 LOCATION: 79..447
 OTHER INFORMATION: /product= "heavy chain variable
 OTHER INFORMATION: region C21-H1"
 US-08-127-721A-11

Query Match	3.1%	Score 39.4	DB 3	Length 468
Best Local Similarity	51.4%	Pred. No. 0.019		
Matches	91	Conservative	0	Mismatches 86
			Indels	0
			Gaps	0
QY	gcagtgaggcaatccctgcctgcagacgcttccctgcagcagcttccctctgggaagtgatcat	323		
Db	405 gtagccgtagattgcctgcgcgctgaaatggctgaacctggcgcagatgtaacggcgctgac	346		
QY	324 caacatgttcagatctcggctgcggccacatgagtgcttcttgctgctgatactcagtgga	383		
Db	345 ctccgtgtgtcagagcgtctcagctccatgtagcgcgtgttgatgctgctgcgcgcgtgaa	286		
QY	384 tctgctctggggaagtcacaaacttgcgcagtttgttgtatgtgtcgtgagtgagt	440		
Db	285 gctggcccttgaccttgaaacttctccttttagttgtgtggaagtcgcgcggccttat	229		

RESULT 8
US-08-485-245A-11/c
Sequence 11, Application US/08485245A
Patent No. 6072035
GENERAL INFORMATION:
APPLICANT: Hardman, No. 6072035man
APPLICANT: Kolbinger, Frank
APPLICANT: Saldamnia, Jose

```

; TITLE OF INVENTION: Reshaped monoclonal antibodies against an
; TITLE OF INVENTION: Immunoglobulin isotype
; NUMBER OF SEQUENCES: 55
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 6072035artis Patent Department
; STREET: 59 Route 10
; CITY: East Hanover
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07936-1080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/485,246A
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/127,721
; FILING DATE: 27-SEPTEMBER-1993
; APPLICATION NUMBER: US 07/952,802
; FILING DATE: 25-SEPTEMBER-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 6072035ak, Henry P.
; REGISTRATION NUMBER: 33,200
; REFERENCE/DOCKET NUMBER: 4-19276/A/P2/CIP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (908) 277-5110
; TELEFAX: (908) 277-4306
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 468 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 22..447
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 79..447
; OTHER INFORMATION: /Product="heavy chain variable
; OTHER INFORMATION: region C21-H1"
; US-08-485-246A-11

Query Match 3.1%; Score 39.4; DB 3; Length 468;
Best Local Similarity 51.4%; Pred. No. 0.019;
Matches 91; Conservative 0; Mismatches 86; Indels 0; Gaps 0;

QY 264 gcaatgggaacatcctcgtgaagcgtctcctgagccagtccttcctggaagtgatc 323
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 405 gTAGTGTGATGTCCTCCCTGCAAGTGCTGAACCTGGCGCACTAGTAACGCGGTGTC 346
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 324 ccaactgttcagatctcgtcgtccacatgagtgcttgctggtcgtgatctcagtgga 383
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 345 CTGGCTGGGTCAAGGCTCTAGCTCCATGTAGGCGGTGCTGCTGTGCGCGGTGA 286
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 384 tgcctgtctggggaagtcacacttgagcagtggtgtgtatgtgtcgtgtgaagt 440
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 285 GGAGGCTTGCCCTTCAACTTCCTGTTGATGTTGCTGTGCAAGTGCGCGGCTGAT 229

RESULT 9
US-07-781-034-4
; Sequence 4, Application US/07781034
; Patent No. 5442050
; GENERAL INFORMATION:
; APPLICANT: Fishman, Jay A.
; TITLE OF INVENTION: Molecular Cloning of Antigens Shared By
```

```

; TITLE OF INVENTION: Rat- and Human-Derived Pneumocystis Carinii
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Millitia Drive
; CITY: Lexington
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02173
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/781,034
; FILING DATE: 19911018
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Granahan, Patricia
; REGISTRATION NUMBER: 32,227
; REFERENCE/DOCKET NUMBER: MGH91-02A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 861-6240
; TELEFAX: (617) 861-9540
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1189 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; US-07-781-034-4

Query Match 3.0%; Score 38; DB 1; Length 1189;
Best Local Similarity 57.6%; Pred. No. 0.09;
Matches 68; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

QY 393 ggggaagtcacactggcgagcttggtgtgatgtgtcgtgtgagtgacagcttagg 452
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 544 GGTGAATGTGACCCAGCATGAGAGATCGCTGCTGATGATGATGATGATGATGCT 603
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 453 caacctgagatggtcatcagtaatatctcaacacagactcaccatgaacatgatg 510
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 604 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAG 661

RESULT 10
PCT-US92-08328-4
; Sequence 4, Application PC/TUS9208328
; GENERAL INFORMATION:
; APPLICANT: Fishman, Jay A.
; TITLE OF INVENTION: Molecular Cloning of Antigens Shared By
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Millitia Drive
; CITY: Lexington
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02173
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/08328
; FILING DATE: 19920930
```


RESULT 11
 US-08-332-766A-1
 Sequence 1, Application US/08332766A
 Patent No. 5843647
 GENERAL INFORMATION:
 APPLICANT: JEFFERYS, Alec J.
 APPLICANT: ARMOUR, John
 TITLE OF INVENTION: SIMPLE TANDEM REPEATS
 NUMBER OF SEQUENCES: 125
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: CUSHMAN DARRY & CUSHMAN, L.L.P.
 STREET: 1100 New York Avenue, N.W.
 CITY: Washington
 STATE: D. C.
 COUNTRY: U.S.A.
 ZIP: 20005-3918
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/332,766A
 FILING DATE: 01-NOV-1994
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: GB 9346052.9
 FILING DATE: 21-DEC-1993
 ATTORNEY/AGENT INFORMATION:
 NAME: BIRD, Donald J.
 REGISTRATION NUMBER: 25,323
 REFERENCE/DOCKET NUMBER: 217111/M94/0434/GB
 TELECOMMUNICATION INFORMATION:

RESULT 12
 US-08-476-17/c
 Sequence 17, Application US/08476176B
 Patent No. 5958708
 GENERAL INFORMATION:
 APPLICANT: Hardman, No. 5958708man
 APPLICANT: Kolbinger, Frank
 APPLICANT: Saldanha, Jose
 TITLE OF INVENTION: Reshaped monoclonal antibodies against an
 TITLE OF INVENTION: immunoglobulin isotype
 NUMBER OF SEQUENCES: 35
 CORRESPONDENCE ADDRESSES:
 ADDRESSEE: No. 5958708artis Patent Department
 STREET: 59 Route 10
 CITY: East Hanover
 STATE: New Jersey
 COUNTRY: USA
 ZIP: 07936-1080
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/476.176B
 FILING DATE:
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/127.721
 FILING DATE: 27-SEPTEMBER-1993
 APPLICATION NUMBER: US 07/952.802
 FILING DATE: 25-SEPTEMBER-1992
 ATTORNEY/AGENT INFORMATION:
 NAME: No. 5958708ak, Henry P.
 REGISTRATION NUMBER: 33.200
 REFERENCE/DOCKET NUMBER: 4-19276/A/P2/CIP
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (908) 277-5110
 TELEFAX: (908) 277-4306
 INFORMATION FOR SEQ ID NO: 17:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 467 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: DNA (genomic)
 FEATURE:

THIS PAGE BLANK (ISPTO)

117: gb_est48:*
118: gb_est49:*
119: gb_est50:*
120: gb_est51:*
121: gb_est52:*
122: gb_est53:*
123: gb_est54:*
124: gb_est55:*
125: gb_est56:*
126: gb_est57:*
127: gb_est58:*
128: gb_est59:*
129: gb_est60:*
130: gb_est61:*
131: gb_est62:*
132: gb_est63:*
133: gb_est64:*
134: gb_est65:*
135: gb_est66:*
136: gb_est67:*
137: gb_est68:*
138: gb_est69:*
139: gb_est70:*
140: gb_est71:*
141: gb_est72:*
142: gb_est73:*
143: gb_est74:*
144: gb_est75:*
145: gb_est76:*
146: gb_est77:*
147: gb_est78:*
148: gb_est79:*
149: gb_est80:*
150: gb_est81:*
151: gb_est82:*
152: gb_est83:*
153: gb_est84:*
154: gb_est85:*
155: gb_est86:*
156: gb_est87:*
157: gb_est88:*
158: gb_est89:*
159: gb_est90:*
160: gb_est91:*
161: gb_est92:*
162: gb_est93:*
163: gb_est94:*
164: gb_est95:*
165: gb_est96:*
166: gb_est97:*
167: gb_est98:*
168: gb_est99:*
169: gb_est100:*
170: gb_est101:*
171: gb_est102:*
172: gb_est103:*
173: gb_est104:*
174: gb_est105:*
175: gb_est106:*
176: gb_est107:*
177: gb_est108:*
178: gb_est109:*
179: gb_est110:*
180: gb_est111:*
181: gb_est112:*
182: gb_est113:*
183: gb_est114:*
184: gb_est115:*
185: gb_est116:*
186: gb_est117:*
187: gb_est118:*
188: gb_est119:*
189: gb_est120:*

190: gb_est121:*
191: gb_est122:*
192: gb_est123:*
193: gb_est124:*
194: gb_est125:*
195: gb_est126:*
196: gb_est127:*
197: gb_est128:*
198: gb_est129:*
199: gb_est130:*
200: gb_est131:*
201: gb_est132:*
202: gb_est133:*
203: gb_est134:*
204: gb_est135:*
205: gb_est136:*
206: gb_est137:*
207: gb_est138:*
208: gb_est139:*
209: gb_est140:*
210: gb_est141:*
211: gb_est142:*
212: gb_est143:*
213: gb_est144:*
214: gb_est145:*
215: gb_est146:*
216: gb_est147:*
217: gb_est148:*
218: gb_est149:*
219: gb_est150:*
220: gb_est151:*
221: gb_est152:*
222: gb_est153:*
223: gb_est154:*
224: gb_est155:*
225: gb_est156:*
226: gb_est157:*
227: gb_est158:*
228: gb_est159:*
229: gb_est160:*
230: gb_est161:*
231: gb_est162:*
232: gb_est163:*
233: gb_est164:*
234: gb_est165:*
235: gb_est166:*
236: gb_est167:*
237: gb_est168:*
238: gb_est169:*
239: gb_est170:*
240: gb_est171:*
241: gb_est172:*
242: gb_est173:*
243: gb_est174:*
244: gb_est175:*
245: gb_est176:*
246: gb_est177:*
247: gb_est178:*
248: gb_est179:*
249: gb_est180:*
250: gb_est181:*
251: gb_est182:*
252: gb_est183:*
253: gb_est184:*
254: gb_est185:*
255: gb_est186:*
256: gb_est187:*
257: gb_est188:*
258: gb_est189:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

ACCESSION	BF603905	GI:11701985
VERSION	BF603905.1	
KEYWORDS	EST.	
SOURCE		
ORGANISM	Bos taurus	
	cow.	
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovinae; Bos.	
REFERENCE	1 (bases 1 to 499)	
AUTHORS	Smith,T.P.L., Casas,E., Stone,R.T., Heaton,M.P., Grosse,W.M., Bennett,G.A., Fahrenkrug,S.C., Freking,B.A., Kohrer,G.A., Laegreid,W.W. and Keefe,J.W.	
TITLE	Design and use of four pooled tissue normalized cDNA libraries for EST discovery in cattle	
JOURNAL	Unpublished (2000)	
COMMENT	Contact: Smith TPL USDA, ARS, US Meat Animal Research Center PO Box 166, Clay Center, NE 68933-0166, USA Tel: 402 762 4366 Fax: 402 762 4390 Email: smith@email.marc.usda.gov Single pass sequencing. Bases called and alt trimmed with phred v0.980904.e. Vector identified by cross_match with the -minscore 18 and -mismatch 12 options.	
FEATURES	PCR Primers FORWARD: AGGAACAGCATGACCAT BACKWARD: GTTTCACGTCACGACG Plate: 50 row: D column: 2 Seq primer: ATTAGGACACATATAG.	
SOURCE	Location/Qualifiers 1..499 /organism="Bos taurus" /db_xref="taxon:9913" /clone_id="MARC_3BOV" /tissue_type="pooled" /lab_host="DH108" /note="Vector: pCMV SPORT6; site.1: XbaI; site.2: XhoI; library made from pooled tissue from marrow, alveolar macrophage, ovary, fetal semitendinosus muscle, and fetal longissimus muscle."	
BASE COUNT	101 a 134 c 133 g 131 t	
ORIGIN		
Query Match	23.5%: Score 294.2; DB 151; Length 499;	
Best Local Similarity	74.9%: Pred. No. 2.8e-70;	
Matches 368; Conservative	0; Mismatches 123; Indels 0; Gaps 0;	
OY	194	tcaactcgaggttcgcggagacacacgttgagacagtggtgccttcaaccttcacgtcgg 253
Db	1	TCAATCTTCTTTTGGGGAGACATTCCTGAGCAGCATATAGCTTCAATCTCTTCTGCTGG 60
OY	254	cagctgtgtgcagtgaggcaatcctgtcctgagacggttcctgagccagttcccttcggga 313
Db	61	TCTCTGGGGGTGCAGCTGGCAGCCCTGCTGGATGAGCTTCCTCGAACCACTCCCTTAGGA 120
OY	314	agggtgcatcaacaactgttcagatcttggtcggcgacacatagttccttgttggtgcgtga 373
Db	121	AGATGGTATGATATATGTCCAGTATTCAGAGACCAACATGATGCTTACATCTGTGCTGA 180
OY	374	tctcagtagtatgctgtccttgagggaagtgcaacttgcgcgacgtgtgtgtgtgtagtgcctgg 433
Db	181	TCTCGCGGGGTGTGCTGCTGGGGAGGTCAACCTGCTCGACGTGGTGAATTATGAGCGCTGA 240
OY	434	tggaggtgacagctttaggcaacctgagagaagtcatcataagtaatatcttcaacacagact 493
Db	241	TAGAGGTACACACCCTTCAGTGCACAGAGCGTGGTACAGATAAGTACCTCGATATGACAA 300
OY	494	accacatgaacatgtagacatcactacgtgttcgcgagcctatttgggtgtctgtggcct 553
Db	301	TGGATGTAAAGCATGATATATATGACACATCTGTGCGCGCTATTTTGGCGCTGACATGTGCTCT 360
OY	554	ggtgcctgcacaaagccctctaccgagggagacggagataaagatcagaacagacagatac 613

Db	361	GCTGCTCCGGAGACGCTTGCCACGCCGATCAGAGGACAAGAATTCAGCACACGACC	420
Oy	614	ccagttgtctgcacatgctggcgccctctctgttggaagtgttctgccaagttcaact	673
Db	421	CCAGTTTCTTCACCATGCTGGGCACCCCTTCTGTGGAATTCTCGCCGAGTTCAACT	480
Oy	674	ctgctgctgct 684	
Db	481	CTGCTCTGCTG 491	
RESULT	3	Bf602079	
LOCUS		Bf602079	
DEFINITION		267224 MARC_3BOV Bos taurus CDNA 5', mRNA sequence.	13-DEC-2000
ACCESSION		Bf602079	
VERSION		Bf602079.1	GI:11699302
KEYWORDS		EST.	
SOURCE		cow.	
ORGANISM		Bos taurus	
		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
		Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;	
		Bovidae; Bovinae; Bos.	
REFERENCE		1 (bases 1 to 488)	
AUTHORS		Smith,T.P.L., Casas,E., Stone,R.T., Heaton,M.P., Grosse,W.M.,	
		Bennett,G.A., Fahrenkrug,S.C., Freking,B.A., Rohrer,G.A., Laegreid	
		,W.W. and Keeler,J.W.	
TITLE		Design and use of four pooled tissue normalized cDNA libraries for	
		EST discovery in cattle	
JOURNAL		Unpublished (2000)	
COMMENT		Contact: Smith TPL	
		USDA, ARS, US Meat Animal Research Center	
		PO Box 166, Clay Center, NE 68933-0166, USA	
		Tel.: 402 762 4366	
		Fax: 402 762 4390	
		Email: smith@meatlab.marc.usda.gov	
		Single pass sequencing. Bases called and alt_trimmed with phred	
		v0.980904.e. Vector identified by cross_match with the -mnscore 18	
		and -minmatch 12 options.	
		PCR Primers	
		FORWARD: AGGAAACAGCATATGACCAT	
		BACKWARD: GTTTCACGTCACGACG	
		Plate: 43	
		Row: E	
		Column: 5	
		Seq primer: ATTAGTGACACTATAG.	
FEATURES		Location/Qualifiers	
SOURCE		1..488	
		/organism="Bos taurus"	
		/db_xref="taxon:9913"	
		/clone_lib="MARC_3BOV"	
		/tissue_type="pooled"	
		/lab_host="DH10B"	
		/note="Vector: pCMV SPORT6; Site_1: XbaI; Site_2: XhoI;	
		library made from pooled tissue from marrow, alveolar	
		macrophage, ovary, fetal semitendonsus muscle, and fetal	
		longissimus muscle."	
BASE COUNT		97 a 137 c 135 g 119 t	
ORIGIN			
Query Match		23.0%; Score 288.2; DB 151; Length 488;	
Best Local Similarity		75.7%; Pred. No. 1.3e-68;	
Matches		370; Conservative 0; Mismatches 118; Indels 1; Gaps 1;	
Oy	604	gcaacgatccccagttgtctgcacatgctggcgccctctctctgttggaagtgttcgccca	663
Db	1	GCAACGACCCCCAGTTTGTTCACCATGCTGGGCACCCCCTTCTGTGGAATTCTCGCCG	60
Oy	664	agttcacaactctgacctctgagaagaatcgaaagaagaatgcccgtgttcaaac	723
Db	61	AGTTTCAACTCTGCTCTCTCGAAGACTTTCAGATGAAA-GAACATGCGCCGTGTCAACAC	119
Oy	724	tactatgctgltagcaagtcagcgtgtgltagcaagcatctcaaggtaatctcttgcgtaccce	783

Query Match	21.6%	Score 270.4	DB 187	Length 416
Best Local Similarity	97.2%	Pred. No. 9,8e-64		
Matches 315	Conservative 0	Mismatches 5	Indels 4	Gaps 4
Db	1	atgagcttaagtaaccggagctgtccgagctgtccctctgggacctgaactg 60		
QY	1	atgagcttaagtaaccggagctgtccgagctgtccctctgggacctgaactg 60		
Db	46	ATGAGCTCTAAGTACCGCGCGCTGTCTCNGCGTCCGCCCTCTGTGGCCCTTAACACTN 105		
QY	61	gaagcactctccatctctccctctctatttttttaaccactatgaagctctcttaagat 120		
Db	106	GAACAGCTCTCATTTCTTCTTATTTTATTTTACCCTATGAGCTCTCTTAAGAGAT 165		
QY	121	caaaaggagctcgctggacatcctacaagtgtggccaagaacctgacgctgataggcgccatt 180		
Db	166	CAAAAGGGGCTCGTGGCATCTATCAAGTTGGCCAAGATCGACCGTGATGGCGGCATT 225		
QY	181	ggcttggagcttcctcaactcgaagttccggagaaacagcttggagcagatgtgaccttaac 240		
Db	226	GGCTTGGGCTTCACCTCGAGTTCCGGAGACACACTGAGCAGTGTGGCTTCAAC 285		
QY	241	ctcttcatagct-ggcgctctgtgtgtgaat-gggaatcctgct-ggagagctctctggc 297		
Db	286	CTCTTCATGCTGTGGGCGCTTNGTGTGCAAGTGGGCAATCTCTGTGGAGCGCTTCTGAGC 345		
QY	298	cag-ttccctctctgggaagtgtgt 320		
Db	346	CAGTTTCCCTTNTGGGAAGTGTGT 369		
RESULT 5				
LOCUS T84327				
DEFINITION T84327 343 bp mRNA EST 16-MAR-1995				
Y337405.r1 Soares fetal liver spleen INFILS Homo sapiens cDNA clone IMAGE:110384 5' similar to gb:X54534 BLOOD GROUP RH(D) POLYPEPTIDE (HUMAN) ; mRNA sequence.				
T84327				
ACCESSION T84327.1 GI:712615				
VERSION T84327.1				
KEYWORDS EST.				
SOURCE human.				
ORGANISM Homo sapiens				
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.				
REFERENCE 1 (bases 1 to 343)				
AUTHORS Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M., Holman, M., Hulman, M., Kucaba, T., Le, M., Lennon, G., Mattr, M., Parsons, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, F., Tivaskis, E., Waterston, R., Williamson, A., Wohlmann, P. and Wilson, R.				
COMMENT The Washu-Merck EST Project				
Unpublished (1995)				
Contact: Wilson RK				
Washington University School of Medicine				
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108				
Tel: 314 286 1800				
Fax: 314 286 1810				
Email: est@watson.wustl.edu				
Insert Size: 1034				
High quality sequence stops: 295				
Source: IMAGE Consortium, LNLN				
This clone is available royalty-free through LNLN ; contact the				
IMAGE Consortium (info@image.lnl.n.gov) for further information.				
Insert Length: 1034 Std Error: 0.00				
Seq primer: M13RP1				
High quality sequence stop: 295.				

Query Match	18.9%	Score 236.8	DB 4	Length 491
Best Local Similarity	71.4%	Pred. No. 1.9e-54		
Matches	354	Conservative	0	Mismatches 137; Indels 5; Gaps
OY	16	ccgcggctctgcgcgcgtcgtccctctctgcccctcaacactgaaagcagctcatt	75	
Db	1	CCAGGTCCTCCCGCTGCTGCTCCCTGTG-GGCTTGTGAGACAGACAGCTTTTAA	59	
OY	76	ctctctctctatctttttaccacatagaagcttccttaagatgaataaaggcttcg	135	
Db	60	CTCCTCTCTCTTTTTCATCCCCACAGACACAGCCAGGTGATGCACA--GGTTCAG	116	
OY	136	gcactctcaactgcttgcgcgaacatctaccgcttgaatgcgcgcattgcttggtcttc	195	
Db	117	GGGAGCTATCAAGTCTCCCGGAAATTTGACCCTCATGTGCACAG--TTGGGTTGGGCTTCTG	175	
OY	196	aacctcgagttcccgagagaacaagctltgaaagctgtgacctcaacctctcatgctgcg	255	
Db	176	TCTCTTCTCTTTTGGAGACACAGCTGGAGCGTGTGGCTTTCACCTCTTCATATGGGCC	235	
OY	256	cttggtgtgcagttggcgaatcctgctgtgaacgyccttcccttgagccaagtctcccttcggaag	315	
Db	236	CTCGGGGTCGAGGGGACAATCTTGCTGGACCATTTCTTG66CCAGGTCCTCCAAATGGAAC	295	
OY	316	gtgtgctcatcaactgcttcagctatctgcgcttgcgcacatgagctcttgctgtgcgtgcg	375	

```

/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="H3134C09"
/clone_1lb="NIA Mouse 15K cDNA Clone Set"
/sex="Clones arrayed from a variety of cDNA libraries"
/dev_stage="Clones arrayed from a variety of cDNA libraries"
/lab_host="DH10B"
/notes="Vector: pSPORT1; Site_1: SalI; Site_2: NotI. This clone is among a rearranged set of 15,247 clones from 11 embryo cDNA libraries (including preimplantation stage embryos from unfertilized egg to blastocyst, embryonic part of E7.5 embryos, extraembryonic part of E7.5 embryos , and E12.5 female mesonephros/gonad) and one newborn ovary cDNA library. Average insert size 1.5 kb. All source libraries are cloned unidirectionally with Oligo(dT )-Not primers. References include: (1) Genome-wide expression profiling of mid-gestation placenta and embryo using a 15,000 mouse developmental cDNA microarray, 2000, Proc. Natl. Acad. Sci. U S A, 97: 9127-9132; (2) Large-scale cDNA analysis reveals phased gene expression patterns during preimplantation mouse development, 2000, Development, 127: 1737-1749; (3) Genome-wide mapping of unselected transcripts from extraembryonic tissue of 7.5-day mouse embryos reveals enrichment in the X-chromosome and under-representation on the X chromosome, 1998, Hum

```


VERSION	T89255.1	GI:717768
KEYWORDS	EST.	
SOURCE	human.	
ORGANISM	Homo sapiens	
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.	
REFERENCE	1 (bases 1 to 368)	
AUTHORS	Hillier,L., Clark,N., Dubugue,T., Elliston,K., Hawkins,M., Holman	
	,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J.,	
	Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevasakis,E., Waterston	
	,R., Williamson,A., Wollmann,P. and Wilson,R.	
	The WashU-Merck EST Project	
TITLE	Unpublished (1995)	
JOURNAL	Contact: Wilson RK	
COMMENT	Washington University School of Medicine	
	444 Forest Park Parkway, Box 8501, St. Louis, MO 63108	
	Tel: 314 286 1800	
	Fax: 314 286 1810	
	Email: est@watson.wustl.edu	
	Insert Size: 1034	
	High quality sequence stops: 297 Source: IMAGE Consortium, LLNL	
	This clone is available royalty-free through LLNL ; contact the	
	IMAGE Consortium (info@image.llnl.gov) for further information.	
	Insert Length: 1034 Std Error: 0.00	
	Seq primer: -21m13	
FEATURES	High quality sequence stop: 297.	
source	Location/Qualifiers	
	1..368	/organism="Homo sapiens"
	/db_xref="GDB:466001"	
	/db_xref="taxon:9606"	
	/clone="IMAGE:110384"	
	/clone_1fb="Soares fetal liver spleen INFLS"	
	/sex="male"	
	/dev_strage="20 week-post conception fetus"	
	/lab_host="DH10B (ampicillin resistant)"	
	/note="Organ: Liver and Splein; Vector: pTR73D (Pharmacia)	
	with a modified polylinker; Site_1: Pac I; Site_2: Eco RI;	
	1st strand cDNA was primed with a Pac I - oligo(dT) primer	
	[5' AACTCGAAGATTAATTAAAGATCTTTTGTGGTTTTTTT 3'] ,	
	double-stranded cDNA was ligated to Eco RI adaptors	
	(Pharmacia), digested with Pac I and cloned into the Pac I	
	and Eco RI sites of the modified pTR73 vector. Library	
	went through one round of normalization. Library	
	constructed by Bento Soares and M.Fatima Bonaldo."	
BASE COUNT	101 a 79 c 81 g 102 t 5 others	
ORIGIN		
Query Match	15.5%; Score 194.6; DB 189; Length 368;	
Best Local Similarity	93.9%; Pred. No. 7.8e-43;	
Matches 200; Conservative	0; Mismatches 13; Indels 0; Gaps 0;	
OY	1042 gtgctgataccgtgcggcagcgcaatgcatgatggtccacaggctccctcagcatt	1101
Db	359 GGCGCATCATTACGTGTCGAACGGCAATGGCAATGGCTGCCAGGTCCTCCACGAT	300
OY	1102 ggaggaccactggtgcgatcgtagtacctcgaagtctgctcttcacagttgttc	1161
Db	239 GGGAGCATCAGCTTGCGCATCGTAGTACCTCCACGCTCGGCTCTCTGCACAGTTGGCTC	240
OY	1162 cttaatctaaatatggaagcacccatagaagcttaaatatttgatgaccagtttc	1221
Db	239 CTAAATCTCAAAATATGGAAGACACCTCATGTGGCTTAATAATTTTGATGACCAGATTTC	180
OY	1222 tggaagttccctcatttggctgtgtgatttaa	1254
Db	179 TGGAA GTTCCATTTGGCTGTTGGATTTTAA	147
RESULT	11	
AM553137/c		
LOCUS	AME53137	645 bp mRNA EST 31-AUG-2000

DEFINITION	L0222E01.3 Mouse Newborn Ovary cDNA Library Mus musculus cDNA clone			
ACCESSION	L0222E01.3', mRNA sequence.			
VERSION	AW53137			
KEYWORDS	AW53137.1 GI:7198560			
SOURCE	EST.			
ORGANISM	Mus musculus			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
AUTHORS	1 (bases 1 to 645) Tanaka,T.S., Jaradat,S.A., Lim,M.K., Kargul,G.J., Wang,X., Grabovac,M.J., Pantano,S., Sano,Y., Piao,Y., Nagaraja,R., Doi,H., Wood,W.H.III, Becker,K.G. and Ko,M.S.H.			
TITLE	Genome-wide expression profiling of mid-gestation placenta and embryo using a 15,000 mouse developmental cDNA microarray			
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 97 (16), 9127-9132 (2000)			
MEDLINE	20381348			
COMMENT	Contact: George J. Kargul Laboratory of Genetics National Institute on Aging/National Institutes of Health 333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA Email: cdna@jgsum.grc.nia.nih.gov Plate: L0222 row: E column: 01 Seq primer: -21M13 Forward High quality sequence stop: 645 POLYA=yes.			
FEATURES	location/Qualifiers			
source	1..645			
	/organism="Mus musculus"			
	/strain="C57BL/6J"			
	/db_xref="taxon:10090"			
	/clone="L0222E01"			
	/clone_lib="Mouse Newborn Ovary cDNA Library"			
	/sex="female"			
	/dev_stage="Newborn Ovary"			
	/lab_host="DH10B"			
	/note="Vector: pSPORT1 (Gibco/BRL Life Technology); Site_1: SalI; Site_2: NotI; Total RNAs were extracted from 7 Newborn Ovary. The double-stranded cDNA was synthesized by Gibco's kit with an oligo(dt) primer [NotI primer-adapter from GibcoBRL] [5'-pGACTGTGTCCTCAATGCGAGCGGCCGCCCTTTT-TTTT-3'] from 2.5ug of total RNA. The double-stranded cDNAs were treated with T4 DNA polymerase and purified by ethanol-precipitation. The cDNAs were ligated to lone-linker lI-Sal3 (include SalI sequence). The cDNAs were purified by phenol/chloroform and separated from free linkers by Centricon 100. Then, cDNAs were amplified by long-range high fidelity PCR using Takara's Ex Taq polymerase. Then, the cDNAs were purified by phenol/chloroform and by Centricon 100. The cDNAs were digested with SalI and NotI enzymes. Then, the cDNAs were size selected by Gibco's Size Fractionation Column. The cDNAs were cloned into SalI/NotI site of pSPORT1 plasmid vector. The DH10B E. coli host was transformed with the ligation mixture by chemical method. The library was constructed by Xiaohong Wang and Yulan Piao."			
BASE COUNT	181 a 175 c 157 g 130 t 2 others			
ORIGIN				
Query Match	15.0%; Score 188; DB 117; Length 645;			
Best Local Similarity	64.6%; Pred. No. 6.2e-41;			
Matches 297; Conservative	0; Mismatches 157; Indels 6; Gaps 1;			
QY	795 cagcaagactatgtcagcacagtgcggtgttggcagaagagcgtgctgtggttacctcgtg 854			
DB	645 CAACATGTTGCATCTACATCCACATGCAGTGTGGCGAGGAGCGCGCGTGGCGGCCCGG 586			
QY	855 tcaactgtatcccttcttcgtgtgtcgtccatggtgtcgtgtgtgtgtgtgtgtgtgtc 914			
DB	585 TTGGCTGATTTCTTCACCTTGATTTCCATGATGTGCTGGCGCCCTCAATGATGGGTTGATCTC 526			

QY	915	cgctggggaaacgaagcagcactgcgcggggtcttctaacagatgcctggattccccaag	974
Db	525	CATCTGGGGACCAAGTGTCCACGGGGGTTTGACACACATCTCTGANA-----ACTC	472
QY	975	ctccacatcagtgctacaacttcagcttgcctggctcgtcttggagagatcattcatgt	1034
Db	471	CAGTGGGATCCACACACACTTTCGGCTTCGGGTCTGCTGGGAGCACTTACTACTACTG	412
QY	1035	gctcgtcgtctgtatcacgcctggagcgcgcgaatgycatgatctgagctcagctccct	1094
Db	411	CCTTCACGATAGTGACACAGGCCCAAGTCTCTGGATCTCTGGATCATACCCACAGCGTAC	352
QY	1095	cagcatctgggggaactcagccttggccatcgtatagcttcaacgtctcgtctctcgcag	1154
Db	351	TCACATTGGGGCTCTNNATGTTCCCTGTGGCGGAGGGTATGATGACAGCACTCTCAAGG	292
QY	1155	tttgctccctaaactcttaaaatcagaaagacctctcagagctctaaatatttgcagaca	1214
Db	291	TTGTCTCTTAAGTGTCAAGATGTGGAGGGCTCCCGCTGGGCCCAAGTATTGATGATCA	232
QY	1215	agtttcttgaaagtctcccatcttgctcgttgcattttaa	1254
Db	231	GACTTCTGGGAGTTCCCAACTTTGGGGGTGGATTTTTAA	192

LOCUS	DEFINITION	AA008551	RESULT 12
AA008551	370 bp mRNA	EST	25-JUL-1996
cg86a09.1	Soares mouse embryo	NbME13.5 14.5 Mus musculus	CDNA
clone IMAGE:439864	's similar to gb:X55534	BLOOD GROUP RH(D)	
POLYPEPTIDE (HUMAN);	mRNA sequence.		

ACCESSION	AA008551
VERSION	AA008551.1
KEYWORDS	GI:1464518
SOURCE	EST. house mouse.
ORGANISM	Mus musculus

REFERENCE
AUTHORS
Eukaryota: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi,
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 370)
Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.

TITLE	The WashU-HHMI Mouse EST Project
JOURNAL	Unpublished (1996)
COMMENT	Contact: Marra M/Mouse EST Project

WashU-HHMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel.: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LNL, the
IMMGE Consortium (infoelimage.llnl.gov) for further info

Seq primer: ETPrimer
High quality sequence stop: 271.

```

FEATURES
source      Location/Qualifiers
1. .370

```

```

/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:439864"
/clone_lib="Scars mouse embryo NbME13.5.14.5"
/sex="unknown"

```

```

/issue_type="embryo"
/dev_stage="13.5-14.5dpc total fetus"
/lab_host="IDH10b"
/notes="vector: pT73b-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
was primed with a Not I - oligo (dT) primer [5',
TGTTCACCAATCTGAAGTGGAGCGCGCCGGGAAAATTTTTTTTTTTTTTTTTT

```

T 3], on equal amounts of mRNA from 2 13.5dpc and 2 14.5dpc embryos [total RNA provided by Minoru KO, Wayne State Univ., from 2]; double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7m3 vector. Library went through one round of normalization, and was constructed by Bento Soares and M.Fátima Bonaldo.

Query Match	14.8%	Score 186.2	DB 1	Length 370
Best Local Similarity	73.9%	Pred. No. 1.6e-40		
Matches 277; Conservative	0	Mismatches 93	Indels 5	Gaps 3

OY 5 gctcaatcacccgcggtcgttcgcggcgccctctggagcccacaacatcgaag 64
 | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Ddb 1 gccttaattaccacagggtccctccgcctgctgcccccttgt -gcccttgagctacaga 59

QY 65 cagctctcaatctctctctctctatcttttttaaccaataaagcttccttagagatcaaa 124
||||| + ||||||| + ||||||| + ||||||| + ||||||| +
Db 60 CAGCTTTAGTCTCCCTTCTCTTTTTCATCCCCACGACACAGCCAGGTGATCACA 119

QY 125 aggggctcgcgcacatccatacaagtcgccaagaatctgacgcgtgatcgcgccattcgct 184
||| ||| ||||||| || | ||||| ||| ||| |||
Db 120 --GCTTCATGGCGAGCTATCAAGTCCGCCGAATTGTACCCCTCATGGCACCC-TTGCGCT 175

OY 185 tgggcttcctcaacctgagttccgcgagacacagctgtgagcagtgtgccttaacctc 244
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 176 TCGGCTTCCTGTCCCTGGTCCCTTGCGAGACACACAGCTGAGCAGTGTGGCTTCAACCTCT 235

QY 245 tcaagctgtagcgttggttgttgacgtlgygacaatccctgtctgaagcgcttccttaagccaattcc 304
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 236 TCATGTTGGCCCTCGGGGTGACGGGAACAATCTTGCTGCACCAATTTCCTGGGCCAGGTCC 295

QY 305 cttcctgggaaggtgtgcatcacactgttcagtattcgctgccaacatgagtccttgc 364
 |||| | ||| | ||| | ||| | ||| ||| | |||
 Db 296 TCCAATGGAACAAGATCAACAATCTGTCCAGCATCCAGATAGCTAACCATGACACCTTAC 355

QY	365	cggcgcgtgacatcag	379
Db	356	CTGTCTGATCTCAG	370

RESULT 13
AL363070
LOCUS
AL363070
400 bp
EMBL
EMBL
04-APR-2000

DEFINITION	mus musculus cDNA clone K9304B29.5
ACCESSION	U00000.1
VERSION	AL363070.1
KEYWORDS	EST.
SOURCE	house mouse.
ORGANISM	Mus musculus

REFERENCE
AUTHORS

1 (bases 1 to 400)

Eickhoff, H., Schuchhardt, J., Ivanov, I., Meler-Ewert, S., O'Brien, J.,
Eunaryota, Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

Malik, A., Tandon, N., Wojski, E., Rohls, E., Nyasik, L., Reinhardt, R.

TITLE	, Nietfeld, W. and Lehrach, H. Tissue gene expression analysis using arrayed normalized cDNA libraries
JOURNAL	Genome Res. (2000) In press
COMMENT	Contact: MPIG Abt. Lehrach

FEATURES	Max Planck Institut Fuer Molekulare Genetik Innestrasse 73, Berlin, 14195 Germany The cDNA libraries ICRF 522 and 523 were normalized with oligonucleotide fingerprinting, resulting in a unique subset of 5376 cDNA clones.
SOURCE	Location/Qualifiers 1.400

Db 418 TGATCAGGCTTCTGGGAGTTCCCCCACTTGGCTGTCG 455

RESULT 15	LOCUS	DEFINITION
A1632247/c	482 bp	NCI CGAP_G66 Homo sapiens CDNA clone IMAGE:2241355 3'

BASE COUNT	128 a	106 c	95 g	153
ORIGIN				

Query Match	14.28;	Score 178.2;	DB 22;	Length 482;
Best Local Similarity	98.48;	Pred. No. 2.9e-38;		
Matches 180; Conservative	0;	Mismatches 3;	Indels 0;	Gaps 0;

QY	1252	taa	1254
Db	197	TAA	195

Search completed: September 12, 2001, 12:18:45
Job time: 5328 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 12, 2001, 10:49:57 ; Search time 1728.72 Seconds
(without alignments)
11220.185 Million cell updates/sec

Title: US-09-600-714-41
Perfect score: 1254
Sequence: 1 atgagctctaagtaaccgcgcg.....atttgctgttgattttaa 1254

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1344157 seqs, 7733874588 residues

Total number of hits satisfying chosen parameters: 2688314

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl: *
1: gb_da1:*
2: gb_da2:*
3: gb_da3:*
4: gb_in1:*
5: gb_in2:*
6: gb_in3:*
7: gb_om:*
8: gb_ov:*
9: gb_pat1:*
10: gb_pat2:*
11: gb_ph:*
12: gb_pl1:*
13: gb_pl2:*
14: gb_pl3:*
15: gb_pl4:*
16: em_da1:*
17: em_da2:*
18: em_fun:*
19: em_htgo_hum:*
20: em_htgo_inv:*
21: em_htgo_rod:*
22: em_htg_hum1:*
23: em_htg_hum2:*
24: em_htg_hum3:*
25: em_htg_hum4:*
26: em_htg_hum5:*
27: em_htg_hum6:*
28: em_htg_hum7:*
29: em_htg_hum8:*
30: em_htg_inv1:*
31: em_htg_inv2:*
32: em_htg_other:*
33: em_htg_rod:*
34: em_hum1:*
35: em_hum2:*
36: em_hum3:*
37: em_hum4:*
38: em_hum5:*
39: em_hum6:*
40: em_hum7:*
41: em_in:*
42: em_om:*
43: em_or:*

44: em_ov:*
45: em_pat:*
46: em_ph:*
47: em_pl:*
48: em_ro:*
49: em_sts:*
50: em_sy:*
51: em_un:*
52: em_v1:*
53: gb_sts1:*
54: gb_sts2:*
55: gb_sts3:*
56: gb_sy:*
57: gb_un:*
58: gb_v11:*
59: gb_v12:*
60: gb_htg1:*
61: gb_htg2:*
62: gb_htg3:*
63: gb_htg4:*
64: gb_htg5:*
65: gb_htg6:*
66: gb_htg7:*
67: gb_htg8:*
68: gb_htg9:*
69: gb_htg10:*
70: gb_htg11:*
71: gb_htg12:*
72: gb_htg13:*
73: gb_htg14:*
74: gb_htg15:*
75: gb_htg16:*
76: gb_htg17:*
77: gb_htg18:*
78: gb_htg19:*
79: gb_htg20:*
80: gb_htg21:*
81: gb_htg22:*
82: gb_htg23:*
83: gb_htg24:*
84: gb_htg25:*
85: gb_pr1:*
86: gb_pr2:*
87: gb_pr3:*
88: gb_pr4:*
89: gb_pr5:*
90: gb_pr6:*
91: gb_pr7:*
92: gb_pr8:*
93: gb_pr9:*
94: gb_ro1:*
95: gb_ro2:*
96: gb_in4:*
97: gb_pr10:*
98: em_ba3:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1254	100.0	1254	9	AX022514	AX022514 Sequence
2	1252.4	99.9	1254	85	AB018969	AB018969 Homo sapi
3	1252.4	99.9	1354	97	HUMRHDANTI	L08429 Human Rhd b
4	1252.4	99.9	1545	93	HSRRII	X63094 H.sapiens m
5	1252.4	99.9	2790	93	HSRRII	X63097 H.sapiens m
6	1250.8	99.7	1254	85	AB018966	AB018966 Homo sapi
7	1250.8	99.7	1254	85	AB018967	AB018967 Homo sapi
8	1247.6	99.5	1260	97	S78509	S78509 Rhd [human,

9	1246	99.4	1254	85	AB018968	AB018968 Homo sapi
10	1246	99.4	1254	97	S82449	S82449 Homo sapien
11	1242.8	99.1	1458	89	AF187846	AF187846 Homo sapi
12	1239.8	98.9	1251	97	S57971	S57971 Rh polypt
13	1239.6	98.9	1254	85	AB046420	AB046420 Homo sapi
14	1238	98.7	1228	85	AB049754	AB049754 Homo sapi
15	1231.4	98.2	1247	88	AF037626	AF037626 Homo sapi
16	1226.8	97.8	1254	85	AB037270	AB037270 Homo sapi
17	1225.2	97.7	1336	89	AF312679	AF312679 Homo sapi
18	1222	97.4	1330	89	HS297026	HS297026 Homo sapien
19	1220.4	97.3	1254	85	AF177938	AF177938 Homo sapi
20	1199.6	95.7	1328	85	AB030388	AB030388 Homo sapi
21	1199.6	95.7	1328	85	AB049753	AB049753 Homo sapi
22	1191.6	95.0	1350	91	GORRH	GORRH Gorilla gor
23	1188.4	94.8	1328	85	AB018644	AB018644 Homo sapi
24	1186.8	94.6	1350	91	CHPRC	CHPRC Pan troglod
25	1185.2	94.5	1384	10	189819	189819 Sequence 1
26	1185.2	94.5	1384	97	HMRHRA	M34015 Human Rh po
27	1183.2	94.5	1384	97	HSRH30A	X34534 Human Rh po
28	1183.6	94.4	1328	85	AB018645	AB018645 Homo sapi
29	1183.6	94.4	1350	91	CHPRH	L37049 Pan troglod
30	1182	94.3	1466	10	189820	189820 Sequence 2
31	1180.8	94.2	1427	88	AF056965	AF056965 Homo sapi
32	1180.6	94.1	1251	97	S57967	S57967 Rh polypt
33	1180.4	94.1	1350	91	CHPRH	L37048 Pan troglod
34	1169.2	93.2	1350	91	GORRA	L37052 Gorilla gor
35	1162.2	87.9	1270	88	AF012425	AF012425 Pongo pyg
36	1055.6	84.2	1350	97	MACRA	L37054 Macaca fasc
37	1049.2	83.7	1369	97	S70343	S70343 Rh-like (Ma
38	1038	82.8	1266	88	AF012426	AF012426 Papio ham
39	972.8	77.6	1263	88	AF012429	AF012429 Callithr
40	972	77.5	1263	88	AF012428	AF012428 Salimr
41	962.8	76.8	1216	88	AF101479	AF101479 Cebus ape
42	954.2	76.1	1209	88	AF102876	AF102876 Cebus ape
43	952	75.9	1216	88	AF052588	AF052588 Callithr
44	941.2	75.1	951	97	S73913	S73913 Rh D-RhX62
45	939	74.9	966	97	S70174	S70174 Rhd-Rh b100

ALIGNMENTS

RESULT 1						
AX022514	AX022514	1254 bp	DNA	PAT	07-SEP-2000	
LOCUS	Sequence 41 from Patent WO937763.					
DEFINITION	AX022514					
ACCESSION	AX022514					
VERSION	AX022514.1	GI:10046112				
KEYWORDS						
SOURCE	unidentified.					
ORGANISM	unclassified.					
REFERENCE	1 (bases 1 to 1254)					
AUTHORS	Flegel, W.A. and Wagner, F.F.					
TITLE	Novel nucleic acid molecules correlated with the rhesus weak d					
JOURNAL	phenotype					
	Patent: WO 93/7763-A 41 29-JUL-1999;					
	FLEGEL, WILLY A (DE); WAGNER, FRANZ F (DE); DRK BLUTSPENDEDIENST					
	BADEN WUE (DE)					
FEATURES	Location/Qualifiers					
source	1..1254					
	/organism="unidentified"					
	/db_xref="taxon:32644"					
	1..1254					
	/note="unnamed protein product"					
	/codon_start=1					
	/protein_id="CAC07879.1"					
	/db_xref="GI:10046113"					
CDS	/translation="WSSKYPKSVRRCPLMALITLLEALILFFYTHVDASLEDKGL					
	VASYGVODLTVMALIGLGLTSSFRHSWSSVAFNLPMALGLGYMALILGLFSLQFP					
	SGKVVITLFSILATMSALSVLISVDALVGLVLMVYVETALGNLGMISQNP					
	FMIDYHMNMHIYFAAYEGSLVACDKRPIDEGEDDQDRTITSLAMGALPLM					
	FMPSNLSLRPIERKNVFTTYAVAVSVYTAISGSSLAHPQCKISKITYHSAVLA					

Db 901 GCTGGGCTGATCTCCGTCGGGGAGCCAGTACCTGCCGGGTGTGTAAACCGAGTGTCTG 960
Qy 961 gggattcccaagctccatcatatggtccacactcagcttgcgggtctctgctggag 1020
Db 961 GGGATTCCCAAGCTCCATCATATGGCTGACAACTTCAAGCTTGGGGTCTGCTGGAGAG 1020
Qy 1021 atcatctacattgtctgtctgtctgtatccgctcgagccggcaatgagcatattggc 1080
Db 1021 ATCATCTACATTGTGCTGCTGTGTATACCTCGAGCCGGCATATGATATTTGGC 1080
Qy 1081 ttccaggtccctcccaagcattgggaactcagcttggccatcgtagactctcaact 1140
Db 1081 TTCCAGGTCCCTCCAGCATTTGGGAACCTGAGTGGCATATGATATGCTTCACGCTCT 1140
Qy 1141 ggtctctgaaaggttgcctctcaataatcttaataatgaaagacacatcgagcttaa 1200
Db 1141 GGTCTCTGAAAGGTTGCTCTCAATCTTAAATATGGAAGCAACCTCATGAGGCTTAA 1200
Qy 1201 tatttgatgaccaaatttctggaatttccatattgtctgtgatttaa 1254
Db 1201 TATTGTGATGACCAAGTTTCTGGAAGTTTCTCATTTGGCTGTGATTTTAA 1254

RESULT 2
AB018969
LOCUS AB018969 1254 bp mRNA PRI 03-JUL-1999
DEFINITION Homo sapiens RHDY0 mRNA for Rh blood group D antigen (RHD), complete cds.
ACCESSION AB018969.1 GI:5360898
VERSION RHDY0: Rh blood group D antigen (RHD).
KEYWORDS Homo sapiens (isolate:yo) cdna to mRNA.
SOURCE Homo sapiens
ORGANISM Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (stee) Hyodo, H., Ishikawa, Y., Kashiwase, K., Ogawa, A., Watanabe, Y., Taneyama, H., Toyoda, C., Uchikawa, M., Akaza, T. and Fujii, T. Polymorphisms of RHDVa in Japanese
JOURNAL Unpublished (1998)
TITLE 2 (bases 1 to 1254)
AUTHORS Uchikawa, M., Hyodo, H. and Ishikawa, Y.
REFERENCE Direct Submission
JOURNAL Submitted (20-OCT-1998) to the DDBJ/EMBL/Genbank databases.
TITLE Hironobu Hyodo, Japanese Red Cross Central Blood Center, Research; 4-1-31, Hiroo, Shibuya, Tokyo 150-0012, Japan
AUTHORS (E-mail: hyodo@hla.cbc.jrc.or.jp, Tel: +81-3-5485-6009, Fax: +81-3-3406-7892)
JOURNAL Location/Qualifiers

FEATURES
source
1. 1254
/organism="Homo sapiens"
/isolate="yo"
/db_xref="taxon:9606"
1. 1254
/gene="RHDY0"
1. 1254
/note="Rhd variant RHDY0"
/codon_start=1
/product="Rh blood group D antigen (RHD)"
/protein_id="BA82159.1"
/db_xref="GI:5360898"
/translation="MSKYPRSYRRCILPLIALTLEALILLFFTHYDASLEDOKLT
VASYQVGDITVMAALIGLGLISSFRHSHSVAFNIPMLALGVOMLIDGLTSQRP
SGKVVITLISRLATMSALSVLISVDVAVLKVADLVWVLEVVALGLRMVLTINL
FNTDYHNNMMHIVFAAYFGLSVAMCLPKLPGETEDKQDTATIPSLAMGLFLTM
FSPFNSALRSPDIKRNAYFNTYAAVAVVTAISGSSLAHQGLSKTYVSAYVA
GGVAVGSCILSPKLPMLAVLGLVAGLISVGAKYLPGCCNRYLIPHSISMGNFSL
IGLIGETIYVLLVDVNGAGNGMIGVOVLSTIGLSLAIYVLTSLTGLLNLKI
WKAPHEKATYDDVTFKFPPLAVGF"

BASE COUNT 242 a 322 c 350 g 340 t
ORIGIN

Query Match 99.9%: Score 1252.4; DB 85; Length 1254;
Best Local Similarity 99.9%: Pred. No. 0;
Matches 1253; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 atgagctctaagtaaccgagctctgcgcgagctgcgcctccctctggccataactg 60
Db 1 ATGAGCTTAAGTACCCCGGCTGTCCGGGCTGCCCTCCCTCTGGCCCTTAACACTG 60
Qy 61 gaagcagctcattctcctctcctatatttttaaccaactgaagcttcttagaagat 120
Db 61 GAACAGCTCTCATTTCTCTCTTATTTTACCACATATGAGCTTCTTAGAGGAT 120
Qy 121 caaagggcgctgcgcatcctatcaagttgccaagatctgacccggtgagggccatt 180
Db 121 CAAAAGGGCGTCTGGGATCTTATCAAGTTGGCCAAAGATCGAGCTGATGGCCGCAATT 180
Qy 181 ggtctggagcttcacacacagcttccggagacacagctggagcagtgtagcctcaac 240
Db 181 GGTCTGGAGCTTCACACACAGCTTCCGGAGACACAGCTGAGACAGTGTGCTTCAAC 240
Qy 241 ctcttcacgtcgcgcttggtgcaabtggaacatcctcgtgacggtctcctgagccag 300
Db 241 CTCTTCATGCTGCGCGCTGTGTGCAAGTGGCAATCCTGCTGAGCGCTTCTGAGCCAG 300
Qy 301 ttccctctgggaaggtggtgcatcacactgttcaagtttggcctgggccaactatgct 360
Db 301 TTCCCTCTGGGAAGGTGTGTATCACTGTTCAGTATTCGGCTGGCCACCATAGTGTCT 360
Qy 361 ttgcgtgctgcacatcagtgatgacgtctctggggaaggtcaacttggcgcaattggtg 420
Db 361 TTGCGGTGCTGATCTCATGTGATGTCTTGTGGGAAGGTCAACTTGGCCACTTGTGTG 420
Qy 421 gtagtggctggtggaagtgacaagctttagcaactgagatggtcatcagtaataatc 480
Db 421 GTGATGGCTGGTGGAGGTGACAGCTTTAGCAACCTGAGAGTGCATCAGTAATAATC 480
Qy 481 ttcaacacagctccacatgaatgatgacatctacagtggttcgaagcctatttggg 540
Db 481 TTCAACACAGCTCCACATGATGATGATGACATCTACGCTTTCGCGACCTATTTTGGG 540
Qy 541 ctgctgtgagctgtgctgtccaaagcctctaccagaggaaggaagataaagatcag 600
Db 541 CTGCTGTGAGCTGTGTGCTGTCCAAAGCTCTTACCGAGGGAAGAGATAAAGATCAG 600
Qy 601 acagcaacgataccagattgtctgcaatgctggcgccctctctctgtgagatgtctg 660
Db 601 ACAGCAACGATACCCAGATTGTGTGCATGCTGGGCGCCCTCTCTGTGTGATGTCTCG 660
Qy 661 ccaagttcaactctgtctgtcgtgaagaagttcaatcgaaagaagaatgccgtgttaac 720
Db 661 CCAAGTTCAACTCTGTCTGTCTGAGAAGTCAATCAAAAGAAAGAAATGCCGTTTCAAC 720
Qy 721 acctactatgctgtagcaagtcagtcgtgtgacagccatctcaaggtcatctctgtccac 780
Db 721 AACTACTATGCTGTAGCAGTGTAGCGTGTGACACCATCTCAAGGTCTATCTTGGCTCAC 780
Qy 781 ccccaagggaagatcagaagaacttatgtgacagtcgagtcggtgttggcaaggagcgtgct 840
Db 781 CCCCAGGGAAGATCAGAACAACTTATGTGCACAGTGGCGGTGTGCAGAGCGGTGGCT 840
Qy 841 gtgggtacctgtgtcaactctatcccttctccggcgtctgacatgagtggtggtctgtg 900
Db 841 GTGGGTACCTGTGTCACTGTATCCCTTCTCGTGGCTTGCATAGTGTGCGTGTGTG 900
Qy 901 gctggagctgaactcgcgtcgtggggaagcaagtaactgcccgggtgttctaaccgagctgtg 960
Db 901 GCTGGAGCTGATCTCGTGGGGAGCCAGATCTGCCGGGCTTTGTAAACGAGTGTCTG 960
Qy 961 gggattcccaagctccatcatatggtgtacaaactcagcttgcgggtctgtgtgagag 1020
Db 961 GGGATTCCCAAGCTCCATCATGTGGCTACAACTTCAAGCTTGTGGGTCTGTGGAGAG 1020

QY	1021	atccatctacattgctgcgtctggtctgtctgtatctgaaccctgcgaagccggcaatgcatgtatgctgc	1080
Db	1021	ATCATCTACATTGTCTCTGCTGTGCTTGAATACCGTCGAGCCGGCAATGGCATGTATGGC	1080
QY	1081	ttccaggttcctctctaagatttggggaaactccaacttggccatctgtatagctcttcacgtct	1140
Db	1081	ATTCAGAGTCTCTTACGACATTTGGGAACCTACACCTTGGCCATGTATAGCTTCACAGCT	1140
QY	1141	ggtctccctacaggttgcctcctaactcttaaatatgaaagaccctcatgagctaaa	1200
Db	1141	GGTCTCCCTGACAGGTTTGCTCTTAATATCTTAATAATATGAAAGACACTCATGAGCTAAA	1200
QY	1201	tatttgatgaccaagtttcttcgaaagtttccctcatgttgctcgttggatlttaa	1254
Db	1201	TATTTGATGACCAAGTTTCTTGAAAGTTTCTCTCATTTGGCTGTGATTTTAA	1254
RESULT	3		
HOMRHDANTI			
LOCUS	HOMRHDANTI	1354 bp	mRNA
DEFINITION	Human Rhd blood group antigen mRNA, complete cds.		24-AUG-1993
ACCESSION	L08429		
VERSION	L08429.1	GI:337390	
KEYWORDS	Rh blood group; Rhd blood group; antigen; blood group antigen.		
SOURCE	Homo sapiens (library: HLI058b (from Clontech)) Bone marrow CDNA to mRNA.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
AUTHORS	1 (bases 1 to 1354) Arce,M.A., Thompson,E.S., Wagner,S., Coyne,K.E., Ferdman,B.A. and Lublin,D.M.		
TITLE	Molecular cloning of Rhd cDNA derived from a gene present in Rhd-positive, but not Rhd-negative individuals		
JOURNAL	Blood 82 (2), 651-655 (1993)		
MEDLINE	93320449		
FEATURES			
source	Location/Qualifiers		
	1..1354		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	1..1254		
	/standard_name="RHD"		
	/citation-1[1]		
	/codon_start=1		
	/evidence-experimental		
	/product="Rhd blood group antigen"		
	/protein_id="AAA02679.1"		
	/db_xref="GI:337391"		
	/translation="MSKRYPSVRRCLEPLNALILEALLILEFFTHYDASLEDEQKSLVAVSVOGDLTVMAIGLGLTSSPRRHSVAFNLFMLAQVMAILDGFLSOPPSGKVLTLFSLRTMSALSVLISVDVLAIGVLAQVLAIVTALGNLRATVSNINFTYDHHMMHMLIYEAAFEFVSVMACPKPLEPSTEDKDDQATIPSLSAMGALFLMMNFWPENSALMSPIIERKNAVNTYVAASVVTALISSSLAHPOKISKTYVHSAYVLAGVAVGTSCHLIPSPMLAMVGLVAGLISVGGAKYLRDCCRVGLIHSISIMGNSFLGLGLGEIYLVLDVTVGAGNGMIGVOVLSTIGELSLAVIALVLTSGLTGLLNLNTRIKRAPEAAVYFDQVFWKPHLAYPE"		
BASE COUNT	278 a 346 c 369 g 361 t		
ORIGIN			
Query Match	99.9%	Score 1252.4;	DB 97; Length 1354;
Best Local Similarity	99.9%	Pred. No. 0;	
Matches 1253; Conservative	0;	Mismatches 1;	Indels 0; Gaps 0;
QY	1	atgaactctaaatgaccggagtcgtctccggcgctgcctgcgcctcttgaggcctaactg	60
Db	1	ATGAGCTCTAACTACCCCGGGGTCTCCGGCGCTCCGCCCTCTGGGCCCTAACACTG	60
QY	61	gaagcagctctatctctctctatattttttaccacactatgagcttctcttagagat	120
Db	61	GAAAGAGCTCTATTTCTCTCTTCTATTTTTTTTTACCACCACTATGAGCGCTTCCCTTAGAGAT	120
QY	121	cazaaaggagctcgtyggcatctataaaglttggccaagatctgacgtgatgagcgagcatt	180

Db	121	CAAAAGGGGCTCGTGGCATCTTAAAGTTGGCGAAGACTGACCCTGATGGCGGCATTT	180
OY	181	ggcttggagcttcccacccttgagtttccggaaagacacagcttgagagagtgagctctaac	240
Db	181	GGCTTGGGGCTTCCACCTCGAGTTCCGGAGACACAGCTGGAGGCAgTGTGGCTTTCAAC	240
OY	241	ctctcaccagctcagagcttggctgagctggagacatccctgcagacggtcttccctgagca	300
Db	241	CTCTTTCANAGCTGGGGCTTGGTGTGGCAATGGGCAATCTCTGTGACCGCTTCTTAGCCAG	300
OY	301	ttccctctcgggaaggtggtacatcaactggtttcaaglatctcggcttgagcccaatgagct	360
Db	301	TTCCCTTCTGGGAAGGTGGTCATCACACTGTTCAgATTTCGCTGGCCACCATATGATGCT	360
OY	361	ttttcgtgtgctgatactgaatgagatgctgctcttgggaaagtaaaccttggcagagttggt	420
Db	361	TTTGTGGGGTGGATCTTAAGTGAAGTGTCTTGTGGGAAGGTCAACTGGCGAGTTGGTG	420
OY	421	gtgaatgctcgtgagtgagagtgagacagctttagcaacctaggaatggtacatcaatatac	480
Db	421	GTAATGGGTGCTGGTGGAGGTGACAGCTTATAGCAACTGAGATGTGATCACTAGTAAATAC	480
OY	481	ttcaacacagactacacatatgaatgacatctacgtggttcgacgctatlttggg	540
Db	481	TTTCAACACAGACTACCAATGAACATGATGCATCTAGTGTTCGACACTTATTTTGGG	540
OY	541	ctgtgtcttgagctgtgtgctcggccaagaagctctacccgaaaggaggaataaataacag	600
Db	541	CTGTCTTGAGCTGTGTGCTCTGGCCAAGACCTCTTACCCGAGAGAAAGGAGATTAAGATAG	600
OY	601	acagcaacagatacccaagtttgtctgcacatgcttgagcgccctcttcttgagatgtctgg	660
Db	601	ACAGCAACGATACCCCACTTTGTGTGCCATGTCTGGGGCCCTCTGTGTGAGATGTTCGG	660
OY	661	ccaagtttcaactctgtctgtgtgagaaagtcacaaatgaagaagaatgacgtgtgtcaac	720
Db	661	CCAAGTTTCAACTCTGTCTGTCTGTAGAGAGTCACATGGAAGAAAGATATCCCTGTTTCAAC	720
OY	721	acctactatgcttgaagaagtcacagctgtgtgagaagacatatcaaggtgcatcccttgctac	780
Db	721	ACCTACTATGTGTGAAGAGTACAGTGAGTGATACGCCATCTCAAGGGTCATCTTGACTAC	780
OY	781	ccccaagggaagatcagcaagaacttatagtgcaacagtcgagtttctggcagaaagcttgct	840
Db	781	CCCCAAGGAAGATCACCAAGACTATATGTGCACAGTGGGTTGTGGCAGAGCGTGCT	840
OY	841	gttggttacctggtgtcaactgtatccctctcgttggcttgccaatggtgtctgtgttg	900
Db	841	GTTGGGTACCTGCTGTCACTGTGATCCCTTCCCTGGCTTGCATGTGCTGTGGCTTTGTG	900
OY	901	gctggagctgatacttcgttcgggggaagcaaatctacgttcgggggtgttgaacgaagtgctg	960
Db	901	GCTGGGCTGATCTCGTCGGGGGAGCCAAATACTCTCCGGGGTGTGTAAACGAATGCTG	960
OY	961	ggagattccccaacagcttccatcatcagtgaggtctcaaacatttgctgggtctgcttggaag	1020
Db	961	GGGATTCOCCCAAGCTCTCATTCATGTGGGTTAACTTCACTTGCTGGGTCTGTTGGAGAG	1020
OY	1021	atcatctacattgtgtctgtgtgtgtatgataccgttcgagccggcaatgacatgtgtgc	1080
Db	1021	ATCATCTACATTTGTGTGTGTGTGTGATATCCGTGGGAGCCGGCAATGGCATGTATGGC	1080
OY	1081	ttcagagttccctccacagcatctgggaacatcagtttgccatatgttgatagctcttaccgct	1140
Db	1081	TTTCAGAGTTCCTCCACACATTTGGGGAATCTAGCTTGGCCATGTGTATCTTCACTGCT	1140
OY	1141	ggtctccagaaagtttgcctccaactcttaaaatatggaagacacctataagctaa	1200
Db	1141	GGTCTCCTGAAGGTTTGGTCTTAATCTTAAATATGGAAGACACCTCATATAGGCTAAA	1200
OY	1201	taatttgatgacaagatttctctgaaatttccatcaatttgcctgttggatlttaa	1254

Db	i201	TATTTTGTGACCAAGTTTCTTCGGAAGTTTCCTCATTTGGCTGTTGAGATTAA	1254
RESULT	4		
LOCUS	HSHRHI	1545 bp mRNA PRI	07-JUL-1997
DEFINITION	H.sapiens mRNA for rhesus polypeptide (RhII).		
ACCESSION	X63094		
VERSION	X63094.1 GI:36027		
KEYWORDS	isoform; red cell membrane protein; Rh blood group; Rh polypeptide; Rhesus polypeptide.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Cranialia; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.		
AUTHORS	1 (bases 1 to 1545)		
JOURNAL	Colin, Y.		
TITLE	Direct Submision		
REFERENCE	Submitted (04-NOV-1991) Y. Colin, INSERM, U76, INTS, 6, Rue A.Cabanel, 75015 Paris, FRANCE		
AUTHORS	2 (bases 1 to 1545)		
TITLE	Le Van Kim,C., Cherif-zahar,B., Raynal,V., Mourou,I., Lopez,M., Cartton,J.P. and Colin,Y.		
JOURNAL	Multiple Rh messenger RNA isoforms are produced by alternative splicing		
MEDLINE	Blood 80		
REFERENCE	92360855		
AUTHORS	3 (bases 1 to 1545)		
TITLE	Cartton,U.P. and Colin,Y.		
JOURNAL	Molecular cloning and primary structure of the human blood group Rhd polypeptide		
MEDLINE	Proc. Natl. Acad. Sci. U.S.A. 89 (22), 10925-10929 (1992)		
COMMENT	See also X63094-98 & M4015.		
FEATURES	Location/Qualifiers		
SOURCE	1..1545		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/chromosome="1"		
	/tissue_type="bone marrow"		
	/clone="RhII"		
	/rissue_1lb="human bone marrow cDNA"		
	/map="lp35"		
	59..1312		
	/codon_start=1		
	/product="rhesus polypeptide RhII"		
	/protein_id="CAA44808.1"		
	/db_xref="GI:36028"		
	/db_xref="SWISS-PROT:O02161"		
	/translation="MSSKYPSPRSVRCLPIAMLTLEALLILFFPFTHDASLSEPOKGIVASQYOGDLTYMAAIGLGFTTSFRHSMSVAFNPLMALSGQMILIDGLFSOPPSKAYVTTLSTRITNSALSVLSVDVANLVNIAQLVMVLVEVIALGNRVISINIFNDYIHNNMHAIYYEAFVFGSLVAMCKPKLPJETEDKDQVATIPSLSAMGALFLMMFWPENSALRLSPRIERNAVFYAAVASVTAIISGSSLHPQGISIKTYVHASAVLA GGVVNTSCHLISPMLAMVGLVAGLISVGAKLYPGCCRRVIRVIGIPHSSIIMGNFESI LGLILEITIYLIVLDTVGANGMGIMGOVLISIELSLAIYALISLGLILLINKIKI WKAHPEAKYEDDYEFMKFPHLAGF"		
BASE COUNT	335 a 389 c 417 g 404 t		
ORIGIN			
Query Match	99.9%; Score 1252.4; DB 93; Length 1545;		
Best Local Similarity	99.9%; Pred. No. 0;		
Matches 1253; Conservative	0; Mismatches 1; Indels 0; Gaps 0;		
Oy	1 atgaagctctaaatccgagcgtgcgtccggcgctcctgcgccctctgagccctaacaacty 60 		
Db	59 ATGACTCTCAATACCACCGGCTCTCCGCCCTCCTGCCTTGCGCCCTAACACTG 118 		
Oy	61 gaagacgcttcattctcctctctctattttttaccaccaatagacgcttccttiagaagat 120 		
Db	119 GAAGCAGCTCATATCTCTCTCTTCTATTCTTTTATCCCACTATGACGCTTCCCTTAGAGAT 178 		

QY	121	caaaaggggctcgtygcatacctatcaaatftggccaagtctgcacgtygatgcygccaatt	180
Db	179	CAAAAAGGGGGCTCGTGGGACCTCCTATCAATCAAGTGGGCCAAGATCTGACCGTATGGGGCCATT	238
QY	181	ggcttgggcttcctcaacctcgagcttccgggaacacagctgagcagtgtyggccttaac	240
Db	239	GGCTTGGGCTTCCTCACTCCAGTTTCCGGAGACACACAGCTGGAGCAGTGTGGCCCTTCAAC	298
QY	241	ctcttcabtygtycgcttggtygtgcaagtgggaalcctctgttcagcgtctccctgaaccaag	300
Db	299	CTCTTCAGTCTGGCGCTTGGTGTGTCAGTGGGCAATCTGCTGTGACGCGCTTCGTGAGCCAG	358
QY	301	ttccctcttgggaagtyggtatacaacctgtctcaagtatctggctctgycaccaatgatgtct	360
Db	359	TTCCCTTCTGGGAAGTGGGTCTATCACTGTTTCACATATTCGGCTGGCCACATATGAGTCTCT	418
QY	361	ttgtcgtgagctatctcaagtgatgtctctcttgggaagatctcaacttggcagatgtgtg	420
Db	419	TTGTGTGGTGCTGATCTCAGTGGATCTGTCTTTGGGGAAAGTCACTTGGCGCAGTTGGTG	478
QY	421	gtacatgtygtcgtgtygaagtgacagctttaaggcaacctgagatgtylcatcagtaataatc	480
Db	479	GTGATGTGTCGTGGTGGAGGTGACACTTTAGGCAACCTGAGGATGTCATCAGTAATATTC	538
QY	481	ttcaacaacagactaccacaatgaataatgatgacaatctcagtgtctgcagcctatatttggg	540
Db	539	TTCAACAACAAGATACCAATGACATGATGATGACATCTCTGTTTGGTCAGCCATTATTTGGG	598
QY	541	ctgtctctggcctgtgtgctgtgccaagcctctcaaccgggaagaaacggaagataaatacaag	600
Db	599	CTGTCTGTGGCTGTGTGCTGCCAAGCCTCTTACCCGAGGGACGCGAGGATTAATATAG	658
QY	601	acagacaagataccacagattgtctgccaatgctggggcgccctctcttcttgagatgtctcgg	660
Db	659	ACAGCAAGATATCCAGATTTGTCGTCACATGCTGGGAGCCCTCTTGTGTGATGTTGCG	718
QY	661	ccaagttccaactctgtctctgtctgtgaagaatccaatcgaaaggaagaatgycgtgttcaac	720
Db	719	CCAAATTTCAACTCTGCTCTGTGTGAAGTCCAAATCGAAAGAAAMAATGCGGTGTTCAAC	778
QY	721	acctactatgctgtgagcagtaagcgtgtgtgaagacatcaaggtaactcccttgctcaac	780
Db	779	ACCTACTATGCTGTATGCAATGACATGACCTGTGTGACAGGCATCTCAAGGCTCATTCCTTGACTAC	838
QY	781	cccccaaggaagatacaagaagaactaattgtgcacagtgtagtcttggcaggaagcgttgct	840
Db	839	CCCCAAGGGAATCAGACAGACTTATGTGCAACATGGGGGTGTGGCAGAGAGCGGTGCT	898
QY	841	gtgggtacactcggtgtaacctgataccctctccgtgctggtctgcaatgtagtctgggtctgtg	900
Db	899	GTGGGTACTCTGTGTACCTGTATCCCTTCTCCGTGGCTTGGCCATGATGCTGTGGCTTGTG	958
QY	901	gctggagctgatactcgtctgggggaagcaagatcacctgcgcggggttgttgtaaacgaatgtctg	960
Db	959	GCTGGGCTGATCTCGTGTGAGGGGAGCCAAATGACTGTCCAGGGGTGTTGTAAACCAGATGCTG	1018
QY	961	gggaattcccccaacagctccaatcaatctgagcttacaacttcagcttctgtgtctgttggagag	1020
Db	1019	GGGATTTCCCCACAGGTCATCATATGGGGCTATCAACTTCAGTGTCTGTGGTCTGTTGGAGAG	1078
QY	1021	atcatctacatgtgtctgtctgtgtgtatatacagctcggagccggcaatggaatgatgtggc	1080
Db	1079	ATCATCTACATGTGTGTGTGTGTGTGTGTATACCTGTGCGAGCCGGCAATGGCATGTATTTGGC	1138
QY	1081	ttcagagttcctctcaagctttgggaacccaacttggcgaatcgtatgatagtctcaagctct	1140
Db	1139	TTTCAGAGTCTCTCTAAGCAATTTGGGAACTCAACCTTGGCCATGTATATGCTTCAAGTCT	1198
QY	1141	ggtctccctgacaggttctgtcctaaatcttaaatatggaagacccctatgagctaa	1200
Db	1199	GGTCTCTGTACAGGTTTGTCTCTTAATCTTAATAATATGGAAGCAACCTCATGAGGCTAA	1258

QY	1201	tatttggctaccacagttcttcggagagttctcctcattgccttgagctttaa	1254
Db	1259	TATTTTGATGACCAAGTTTCTCGAAGTTTCTCATTTGGCTGTGGATTTTAA	1312
RESULT	5		
LOCUS	HSRHXIII	2790 bp	msRNA
DEFINITION	Hs sapiens mRNA for rhesus polypeptide (RhXIII).		PRI 07-JUL-1997
ACCESSION	X63097.1	GI:36046	
KEYWORDS	isoform; red cell membrane protein; Rh blood group; Rh polypeptide; rhesus polypeptide.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
AUTHORS	1 (bases 1 to 2790)		
TITLE	Colln, Y.		
JOURNAL	Direct Submission		
REFERENCE	Submitted (04-NOV-1991) Y. Colln, INSERM, U76, INTS, 6, Rue A. Cabanef, 75015 Paris, FRANCE		
AUTHORS	2 (bases 1 to 2790)		
TITLE	Le Van Kim, C., Cherif-zahar, B., Raynal, V., Mouro, I., Lopez, M., Carton, J.P. and Colln, Y.		
JOURNAL	Multiple Rh messenger RNA isoforms are produced by alternative splicing		
MEDLINE	Blood 80 (4), 1074-1078 (1992)		
REFERENCE	92360855		
AUTHORS	3 (bases 1 to 2790)		
TITLE	Le van Kim, C., Mouro, I., Cherif-zahar, B., Raynal, V., Cherrier, C., Carton, J.P. and Colln, Y.		
JOURNAL	Molecular cloning and primary structure of the human blood group Rho polypeptide		
MEDLINE	Proc. Natl. Acad. Sci. U.S.A. 89 (22), 10925-10929 (1992)		
COMMENT	93066356		
FEATURES	See also X63094-98 & M34015.		
SOURCE	Location/Qualifiers		
	1..2790		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/chromosome="1"		
	/tissue_type="bone marrow"		
	/clone="RhXIII"		
	/tissue_11b="human bone marrow cDNA"		
	/map="1p35"		
	17..1270		
	/codon_start=1		
	/product="rhesus polypeptide XIII"		
	/protein_id="CAA4811.1"		
	/db_xref="GI:36047"		
	/db_xref="SWISS-PROT:Q02161"		
	/translation="MSKRPYSVRRCPLMALITLALILFEFFTHYDASLEDQKIVASVQODLVVMAIGIGFITSFRRSSVAFNLFMLIGVOMALIDGFLSOFENSGKAVITLFSRIATMSALSVLISVDALGVNLAQIVLVAVTALGNLRVYISNFTNYHHMMHRIIYEAAYFGFVAMCLPKPLPEGEDKDDVATPSISAMGALFLWMTFEPNSALLRSPIERKNVAVNTYVAIVSYTAISGSLAHPOGKISKTVHSANVLAAGVAVGTSCHLIPSPWLAIVGLVAGLISVGGATYLRBCNRVIGIPRISITMGVNFSLGLGLIITVILVLDVVGAGNGMIGFVLLISGLSLATVIALTISGLTLLNLKRWKRAHEAKYEDDQVFPMFPHLAVPF"		
BASE COUNT	657 a	680 c	652 g
ORIGIN	801 t		
Query Match	99.9%	Score 1252.4;	DB 93; Length 2790;
Best Local Similarity	99.9%	Fred. No. 0;	
Matches 1253; Conservative	0;	Mismatches 1;	Indels 0; Gaps 0;
QY	1	atgaagcttaagtaacggaggtcttcggagagttctcctcattgccttgagctttaa	60
Db	17	ATGAGCTCTAAGTACCCCGGCTCTCCGGCGGCTCCTGCCCCCTTG6GCCCTAACACTG	76
QY	61	gaagcagctcctcctcctcattatlttttaccacataagacgcttccttagagat	120

D	b	77	GAACACTCTCATCTCCCTCTTCTTAATTTTAAACCATAAGACGCTTCCTTAAGAT	136
Q	y	121	caaaagggctctgtagcatcctaacaagtggccaagatctgacgtgtagcgcaat	180
D	b	137	CAAAAGGGGCTCGGCATCTCTATCAAGTTGGCCAAGATCTACCGTGTATGGCCCAT	196
Q	y	181	ggcttgggtctcccaactgaatttcggaagcaagcttggaagagtgtggccttaac	240
D	b	197	GGCTTGGGCTCTCTCACTCTGAAGTTCCGGAGCACAGCTGAGGAGATGTGGCTTCAAC	256
Q	y	241	ctcttcaatgctbtagcgtctgtgtcagatgggcaatcctgtctgaacgcttccctgaacg	300
D	b	257	CTCTTTCATGCTGGGCTTGTGTGTGATGGGCAATTCCTCTGAGAGGCTTCTGAGCCAG	316
Q	y	301	ttccctctgggaaggtgtgtcatalcaaacgtttagtattcgtctgtgccaacatgtgct	360
D	b	317	TTTCCTTCTGGGAAGGGTGATCAACAGTTTACATTTGGCTGGCCACCATGTAGTCT	376
Q	y	361	ttgcggtgctcagatcagatgagatgctgtctctbtyggaaagtcacaacttggcagttgtg	420
D	b	377	TTGTGGGTGCTGATCTAGTGGATGCTGTCTTGGGAAGTCAACTGTGGCAGTTGGTGT	436
Q	y	421	gtatagtgctgtgtgtgagatgtgacagcttaaggcaacctgagatgtgtatcaataatc	480
D	b	437	GTATGTGTGTGTGGTGGAGGTGACAGCTTAAAGCAACTGAGATGTGTATCATGTAAATTC	496
Q	y	481	ttcaacacgaactcccaatgaaatcagatgacatctagaigtltcgagacctatttgg	540
D	b	497	TTTCACACAGACTTCCACATGAACATGATGCATCTACGTGTTCGACCCATATTTTGG	556
Q	y	541	ctgtctgttgagctgtgtgtcgtccaaagcctctacccgaggaacgagagataaagatcag	600
D	b	557	CTGTCTGTGGCTGTGTCCTGTGCCAAAGCCCTTACCCGAGGAAACGAGATTAATATCAG	616
Q	y	601	acaacaaagatacccaagtgtgtctgcacagctbtyggagcccttcttbtgagtgtctgg	660
D	b	617	ACACACACGATACCCAAATTTGTCTGCATGTCTGGGCGCCCTTCTTGTGGATCTTCTGG	676
Q	y	661	ccaagtttcaactctgtctctgtcgtcgtgaaagatccaatgaaagaaagatgtccgtgttcaac	720
D	b	677	CCAAGTTTCAACTCTGCTGTGTAAGAGTCCAATGAAAGGAAGATGCCGTGTCAAC	736
Q	y	721	aacctactatgcttagagatgcaagctgtgtgtgagaaagcaatcttaggttcaatccttgtctaac	780
D	b	737	ACCTACTATGCTGTAGAGTCAAGTCAACGTGTGTACGCCATTCAGGGTCAATCTTGTGCTAC	796
Q	y	781	ccccaaaggagaatcagcaagactttagtgcacagtgcggttlttggcaagaaagcgtgtct	840
D	b	797	CCCCAAGGGAAGATCAACAGACTTATGTGTGCACAGTGCAGTGTGGCAGAGCGCTGGCT	856
Q	y	841	gttgagtaacctgtgtcaactgtatccctctccgttggcttgcacagtgctgtggcttctgtg	900
D	b	857	GTGGGTACCTCTGTGTCACTGTATCCCTTCCCTGGCTTCCCATGTGTCTGTGGTCTTGTG	916
Q	y	901	gctbtagtgatctccgttcgggaggaagcaagtaacctccgggtgttgtttaaaccgaatgtgtg	960
D	b	917	GCTGGGCTGATCTCCGTGGGGGAAGCAATCACTGCCGGGTGTGTAAACGAATGCTG	976
Q	y	961	ggagattccccacagcttccatcaatcattgggtctlaaacatttgcgttggctgttggaaag	1020
D	b	977	GGGATTTCCCAACAGCTCAATCATGTGGCTCAAACTTACGTTCTGGGTCTGTTGGAAG	1036
Q	y	1021	atacatcaaatgtgtcgtcgtgtgtgtatagcgtctggagccgggaatgagatgtgttc	1080
D	b	1037	ATCATCTACATATGTGTGCTGTGTGTGATACCGTGTGGAGCCGGCAATGTGATGTATGGC	1096
Q	y	1081	ttccaggtctctctcagcatcttgggaacatcagcttggccatgttatagctctcaacgct	1140
D	b	1097	TTCCAGGTCTCTCTCAAGCATTTGGGGAACATCAGTTTGGCCATGTATAGCTCTACAGCT	1156
Q	y	1141	ggctctcgaacaggttgtctctctaaatctttaaatatggaagcaacctcatgtaggtctaa	1200

Db	1157	GCTCCTCGACAGGTTTCTCTCTAATATCTTAANAATATGAAAGACACCTCATGAGGCTMAA	1216
Qy	1201	tatttgatgaccaaagtttctggaagtttccctaatcttgctgtgttgatttaa	1234
Db	1217	TATTTTGATGACCAAGTTTCTGGAAGTTTCTCTCATTTTGCTGTGGATTGATTTTAA	1270
RESULT	6		
AB018966			
LOCUS	AB018966		
DEFINITION	Homo sapiens Rhbva(FK) mRNA for Rh blood group D antigen (Rhd), complete cds.		
ACCESSION	AB018966		
KEYWORDS	AB018966.1 GI:5360242		
VERSIONS	Rhbva(FK); Rh blood group D antigen (Rhd).		
SOURCE	Homo sapiens (isolate:FK) cDNA to mRNA.		
ORGANISM	Homo sapiens		
REFERENCE	Emura, Y., Matzuo, Chouda, Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
AUTHORS	1 (styles)		
TITLE	Hyodo, H., Ishikawa, Y., Kashiwase, K., Ogawa, A., Watanabe, Y., Tsuneyama, H., Toyoda, C., Uchikawa, M., Akaza, T., Fujii, T.		
JOURNAL	Polymorphisms of Rhbva in Japanese		
REFERENCE	Unpublished (1998)		
AUTHORS	2 (bases 1 to 1254)		
TITLE	Uchikawa, M., Hyodo, H. and Ishikawa, Y.		
JOURNAL	Direct Submission		
FEATURES	Submitted (20-OCT-1998) to the DDBJ/EMBL/GenBank databases.		
source	Hironobu Hyodo, Japanese Red Cross Central Blood Centares, 4-1-31, Hiroo, Shibuya, Tokyo 150-0012, Japan (E-mail:hyodo@hla.cbc.jrc.or.jp, Tel:+81-3-5485-6009, Fax:+81-3-3406-7892)		
gene	Location/Qualifiers		
CDS	1..1254		
	/organism="Homo sapiens"		
	/isolate="FK"		
	/db_xref="taxon:9606"		
	1..1254		
	/gene="Rhbva(FK)"		
	1..1254		
	/gene="Rhbva(FK)"		
	/note="Rhd variant Rhbva(FK)"		
	/codon_start=1		
	/product="Rh blood group D antigen (Rhd)"		
	/protein_id="BA81899.1"		
	/db_xref="GI:5360243"		
	/translation="MSSKPRPSRVSLPIALTLLEALILFFPFTHDASLEDKGIVASVQGDITVMAIGIGFETSPRRHSWSVAFLNMLALGQNALIDGPIESOPFSGKYVTLFSTRLATMSALSVLISVDALGVNLAQLVMVLEVTALDGNLRVTSNIFNTYHMMNHIIYVEAFVGLISVAMCLPKPLPESTEDKQDPAIPLSISAMGALFLWMPVNSNALIRSPIDQRKNAVENTYAAVSATVATISGSLAHPOGKISKTVSHAVLALGGVAVNSCHLIPSPMLAMVGLVAGLISVGAKYLPCCNRVIGIPHSSTMGVNFSLGLIGETIYIVLVDIPVAGNGMIGVQVLLISIELSLAIYIALITSLGLTLLLNKRWKHAHEAKYFEDQVFMKPHLAVGF"		
BASE COUNT	241 a 323 c 351 g 339 t		
ORIGIN			
Query Match	99.7%; Score 1250.8; DB 85; Length 1254;		
Best Local Similarity	99.8%; Pred. No. 0;		
Matches 1252; Conservative	0; Mismatches 2; Indels 0; Gaps 0;		
Qy	1	atgaagctctaaatgaccgagctctgaccgagctgacctgagccctaaacatg	60
Db	1	ATGAGCTCTAAGTACCGGGGCTGTCCGGCGCTGACCTGCGGCCCTTAACACAG	60
Qy	61	gaagcagctcctcctctcctctctatttttttaccacatgacgcttccctagaagat	120
Db	61	GAGCAGCTCTCATTTCTCTCTTCTATTTTATTTTACCACCTATGAGCGCTCTTAAGAGAT	120
Qy	121	caaaagggagctcgatgacatcctataaagtgtgcaagatctgacgtgtatgagcgccatt	180
Db	121	CAAAAGGGGCTCGTGTGACATCCCTATCAAGTTGGCCAAAGATCTACCGTGTATGGCGGCATT	180

OY	181	gsetctggtccctccacctcgaattcccgagagacacgctgagacgctgtgctcttaac	240
Db	181	GGCTTGGGGCTTCCTCCACCTCGAGTTTCGAGACACACAGCTGAGACGTTGGCTTTCAAC	240
OY	241	ctcttcacatgctgagctgtgtgtcagtgaggcaaccctctgtgacggtctctctgacag	300
Db	241	CTCTTTCATGCTGGGGCTTGGGTGTGACGTGGGCAATCCTGCTGGACGGCTTCTGAGGCCAG	300
OY	301	ttccctctgggaaggtggtcaccacacgttcaagtctgcgtctgcccacatlgagtct	360
Db	301	TTTCCCTTCTGGGAAGGTGGTCATCACACGTTCATGATTTGGGTGGCCACCAATGATGTGT	360
OY	361	ttgctcgtgtgcagatcctcaagtatgtctctctgtgggaagtcacacttgcagcaattgtgt	420
Db	361	TTGTGCGGGTGGATCTCAGTGGATGCGTCTTGGGGAAGTCAACTTTGGCGCAGTTTGGTG	420
OY	421	gtgatgtgtcgtgtgtgaggtgtgacagctttagcgaacctgaggaagtgtgtcaagtaatac	480
Db	421	GTGATGTGTGCTGGGGAGGTGACAGCTTTTAGCCCAACCTTAGATGTGTATCATGTAATATC	480
OY	481	ttcaacacagacacacacacagaaacatgatacgaatacctaagtgtctgcaaccatattgg	540
Db	481	TTTCACACACACTACCATATGAAATATGACATATCTACTGTTCGCAACCTATTTTGGG	540
OY	541	ctgtctgtgagcctgtgtgctgtccaagacctctaccccgagggaaacggagataaagaatcaag	600
Db	541	CTGTCTGTGGCTGTGTGCTGCTGCCAAAGCCCTACCCGAGGGAACGGAGATTAAGATCAG	600
OY	601	acaacacagatcacccaatttgtctgcgaatctgtggagccctctctctgtgagatgtctgg	660
Db	601	ACACACACAGATCCCAATTTGTCTGCTGCATCTGGGGCCCTCTCTTGTGTGAGATGTTTGG	660
OY	661	ccaagtttcaactctgtctgtcgtgagaaagtcacaaatcgaaagagaatgtcgtgtgtcaac	720
Db	661	CCAAGTGTCAACTCTGCTGTGCTGAGAGAATGCCAATCCAAAGGAAGAAATCCCGTGTCAAC	720
OY	721	acctacatctgtgaagcaatcgaagcgtgtgtgtacaagccaatcgaagtcatcctctgtgctaac	780
Db	721	ACCTACATCTGTGAAGCAATCGAGCTGTGTACAGCCATCTCAAGGGTCATCTTTGGCTCAC	780
OY	781	ccccaaaggaaagatcagaacgaactatgtgtcagaatgtgcgtgtgttcagagaagcgtgtgct	840
Db	781	CCCCAAGGAAGATCCACAGAACTATGTGTACAGTGTGCGTGTGTGGCAGAGAGCGTGTGCT	840
OY	841	gttggtacctcgtgtgtcaacctgatacccttctcgtgtgctgtgcaatgtgtcgtgtgtgtgtg	900
Db	841	GTGGGTACCTCTGTACACATGATCCCTTTCCTCGGCTTGCCATGTGTGCTGGGTCTTGTG	900
OY	901	gtctggcgtgatactctcgtcgggggagccaagtactctccggggtgtgtgtaacaggagtgtcg	960
Db	901	GCTGGCTGATCTCTCGTGGGGGAGCCAAATACTCTCCGGGGTGTGTAAACGAGAGCTG	960
OY	961	ggagatcccccaagctccatcaatgtgagtaaacattcaagcttgtgtgtgtgtgtgtgtgagag	1020
Db	961	GGGATTCGCCACACTCCATCATATGGGCTACAACTTCAGCTTGCTGGTGTGCTGTGAGAG	1020
OY	1021	atcaatcacatgtgtgcgtcgtgtgtgtgtataccgttcggagccgggaatgtgcatgtattgc	1080
Db	1021	ATCATCTACATTTGTGTGTGTGTGTGTGATACGTCGTGGACCCGGCAATGGCATGATGTGC	1080
OY	1081	ttcaaggtctccctcagaatcttgaggaaactaagcttggccaatgtgtataagctctcaagctct	1140
Db	1081	TTTCAGGTCTCTCAGCATTTGGGGAATCTAGCTTGCCCATCTGTATACCTGTCAAGTCT	1140
OY	1141	ggctcctctgacaggtttgtctcttaaatcttaaaatagtgaagaagcacctcatgaagctaa	1200
Db	1141	GGCTCTCTGACAGGTTTGTCTCTTAATCTTAAATATGTGAAACACACTATGAGGCTAA	1200
OY	1201	tattttatgagcaagttttctgaaagtttcccaattgtgcctgttgattttaa	1254
Db	1201	TATTTTATGACCAAGTTTTCGTGAAATTTCTCATATTTGGCTTTTGGATTTTAA	1254

[illegible]

Db	241	CTTTTCATGCTGGGGCTTTGGTGTGCAATGAGGCAATCTGCTGAGCGCTTCTTGAGCCAG	300
QY	301	ttccctcttggaaagtggtgcatcaacacbtltcaaglatlccgctggcccaatgagtgct	360
Db	301	TTCCCTTCTGGGAAGGtGtGtCATCACACtGTTCAGTAATTCGGGTGGCCACCATGTAGTGCT	360
QY	361	ttgtcgtgtgcgtgcatctcaatgtagatgtcttcttgggaaagtcaacttggcagtgctg	420
Db	361	TTGTGCGGTGtGtGATCTGAGTGGAATGCTGTCCTTGGGGAAGGTCAACTGTGGCGAGTTGGTG	420
QY	421	gtgatgtgtcgtgtgggggtggaacagctttaggaacactttaggatgtgatacaatc	480
Db	421	GTGATGtGtGCTGGtGGAGtGACAGCTTTAGGCAACCTTAGAGtGtGTCTACGTAAATATC	480
QY	481	ttcaacaagaagatacccaatgaaacatgatbacaatctaagtlctcgaagcatttttgg	540
Db	481	TTCAACAACAGACtCTACCAATGATGAATGATGATGATCATCTAGTGTGTGACCATTAATTTGG	540
QY	541	ctgtctctgtgacctgtgtgtccaaagcctctacccggaaggaaacggaagataaagaatcag	600
Db	541	CTGTCTGTGtGtGCTGTGtGCTCCAAAGCCCTCTACCCGAGGGAACCGAGATTAAGATCAG	600
QY	601	aagaacaagaataccacaaatttgtctgcacatgttggagccctcttcttggatgtctcgg	660
Db	601	ACAGACAGATACCAACAGTTGTGCTGCAATGCTGGGGCCCTCTCTTGtGGAATGTTCTGG	660
QY	661	ccaagtttcaacatctgtctgtctgtgagaagtcacatcgaaaggaaagaaatbgtcgtttcac	720
Db	661	CCAAGTTTCAACtCTGCTGTGtGAGAAAGTCCAAATCCAAGGAAGATGCCATGTTTCAAC	720
QY	721	acctactatgtctagcagtcagcgtgtgtgacaagcatctcaagtgatccttggctaac	780
Db	721	ACCTACTATGCTGTGAGCAAGTCAAGGTGTGTGACAGCCATCTCAAGGGTCAATCTTGGCTTAC	780
QY	781	ccccaaaggaaagatcacgaagacttattgtgcacagtcgagtgatttggcaagaagcgctgct	840
Db	781	CCCCAAGGAAGATCACAAAGACTTATGtGCACAGtGcGtGtTGGCAAGAGCCGTGCT	840
QY	841	gtgggtaacctcgtgtgacacacgtatccctcttcacgtgtgtcttgcacatgtgtgtgtcttg	900
Db	841	GTGGGTACCTGtGTCACTGATCCGATCCCTTCGCTGGCTTGCATGtGTGCTGGGTCTGTG	900
QY	901	gctgtggtcgtgatlctcgtgcgggaggaagccaagttacgtgcgggggtgtgttaaccgagtgctg	960
Db	901	GCTGTGtGtGATCTCCGtCGGGGGGAAGCCAAATACCTCCGGGGtGTGTAAACGAGTGTCTG	960
QY	961	gggattcccccaagcttcacatcatgggtctaaacttaagcttgcgtgtgtcttgggaag	1020
Db	961	GGGATTCCCCAACACCTCATCATGtGGGTAAACTTCAAGCTGCTGGGTCTGTTGGACAAG	1020
QY	1021	atcaatcaacatgtgcgtcgtgtgtgtatatacgttcggaagcggaatgacaatgtttgc	1080
Db	1021	ATCATCATACATGtGTGtGtGtGtGtCTTGATACCGTGGAGCCGGCAATGCAATGATTTGGC	1080
QY	1081	ttcagatgcctcctcagcatltgggaaactcagctltggccatgcgtatagctctcacgtct	1140
Db	1081	TTCCAGGtGtCTCTCAACATTTGGGAACCTCAGCTTGGCATGTGATACCTTCACGCTT	1140
QY	1141	ggtctctcgaaaggtttgtctccctaaatcttaaatatgaaagcacctctatgaagctaaa	1200
Db	1141	GGTCTCTCGAAGAGTGTGtGCTTAATCTTTAAATTTGGAAGAAGCACTCATATGAGGCTAAA	1200
QY	1201	tatttgaatgacaagttcttcggaagtttccatcatlttgcctgttgaatttaa	1254
Db	1201	TATTTTATGATCCAAAGTTTCTGGAAGTTTCCATATTGGCTGTGGATTTTAA	1254
RESULT	8		
LOCUS	S78509	578509	1260 bp mRNA PRI 27-SEP-1995
DEFINITION	RHD [human, D--phenotype, erythrocyte, mRNA partial, 1260 nt].		
ACCESSION	S78509		


```

/db_xref="taxon:9606"
/cell_type="reticulocyte"
/ncbi="alloanti-D producing D-positive patient"
1..1254
/feature="RHD"
1..1254
/feature="RHD"
/function="associated with hemolytic disease of the
newborn in a D-positive child"
/note="This sequence comes from Fig. 4; conceptual
translation presented here differs from translation in
publication"
/codon_start=-1
/product="D category IIIC antigen"
/protein_id="AAB37696.2"
/db_xref="GI:9790343"
/translating="MSKIPRSVRRCILPMLALILELFEFTHYDASLEDQKL
VASQYQODLTVAALIGLFTSSFRHSVSAFLMLALFQWAILDGLSOP
SGRVYITLESIRLATMSAMSLISAGAVILKVAOLVYVLEVTALRLMYISNI
FNDYHNMHHIYVFAAYFGLSYAMCLPKPLPSTEDKDOTATIPSLAMGLFLFM
FWPSPSALLRSPILERNNAVENYVAVSVTAISSSIAHOGKISKYVHSAVLA
GVAVSTSCHLIPSPMLAMVLGLVAGLISVGAKEYLPCCNRVLGIPHSIMGNLSL
LGLLGLTIVLLVLDVAGNMGIFQVLLISGLSLAIVALTSLGLLNLKI
MKAPHAKYFDVDFWKFPHLAVGF"
BASE COUNT      240 a      323 c      352 g      339 t
ORIGIN
Query Match      99.4%; Score 1246; DB 97; Length 1254;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1249; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 1 atgagctctaatgacccgggtctgtcggcgctgcgcgccttcggcgcttaacatg 60
    |||||||
DB 1 ATGAGCTCTAATGACCCGGGTCTGTCCGGCGTGCCTGCCCTCTGGCCCTTAACATG 60
QY 61 gaagcagctcctcattctcctctctatcttttaccactatgacgtcttccttagagat 120
    |||||||
DB 61 GAAGCAGCTCATTCCTCTCTATTTTACCCTATGACGCTTCTTAGAGAT 120
QY 121 caaaggggctcgtggacatcctatcaagtggcaagatcctgacgtgtagggcaatt 180
    |||||||
DB 121 CAAAGGGGCTCGTGACATCTATCAAGTTGCCAAGATCTGACCGCTGATGGCCCAT 180
QY 181 ggcctggccttcctcactcagctgaattccggagagacaaactgagcaagtgtgacctg 240
    |||||||
DB 181 GCGTGGGCTTCTCTACCTCGAGTTTCCGGAGACACAGCTGAGAGAGTGGCCCTTCAAC 240
QY 241 ccttcacatgctcgtgtgtgtcagtggtggcaatccctcgtcgaacgtcttcctgagccag 300
    |||||||
DB 241 CTCCTCATGCTGGGCGTGTGTGACAGTGGGCATCCGCTGGACGGCTTCTGAGCCAG 300
QY 301 ttcccttcgggaaggtgtgtatcatcacactgttcagltatcgtggtggccacatgagtgc 360
    |||||||
DB 301 TTCCCTTCTGGGAAGGTGTCATCACTGTTAGTATTCGGCTGGCCACCATGAGTGTCT 360
QY 361 ttgtcgtgctcgtatcgtatgagtcgtctctggggaaggtcaactgtggcagttgtg 420
    |||||||
DB 361 ATGTGGGCTGATCTCCGCGGTGCTCTTGGGGAAGGTCACTTGGCCAGTTGGTG 420
QY 421 gtgatgtcgtcgtggaagtgtacagctttaagcaacctgagatgagtcataatc 480
    |||||||
DB 421 GTGATGTGCTGCTGGAGAGTACAGCTTTAGGCAACCTCGAAGATGCTATCATGTAATTC 480
QY 481 ttcaacacagactacacatgaaatgtagacatctacgtgtgtcgaagcctatttggg 540
    |||||||
DB 481 TTCAACACAGACTACCAATGAACATGATGACATCTACGTTTGGCAGCCTATTTTGGG 540
QY 541 ctgtcgtgctcgtgtgtgtcgtccaaagccttacccggaggaaggaaggaataaagatcag 600
    |||||||
DB 541 CTGTCTGTGCTGCTGCTGCTGCAAGCCTCTACCCGAGGGAACGAGGATTAAGATCAG 600
QY 601 aacgaacagataccagattgtctgcacatgtctggcgccctctctgttgatgtctcgg 660

```

```

|||||
DB 601 ACAGCAAGATACCCAGTTTGTCTGCCATGCTGGCGCCCTTCTGTGATGATCTCG 660
QY 661 ccaagttcaactctgtctgtcgtgaagaatccaatcgaaagaagatgcgttctaac 720
    |||||||
DB 661 CCAAGTTCAACTCTGTCTGTCTGAGAAGTCCAAATCGAAGAAAGAAATGCCGTCTTCAAC 720
QY 721 accacatgtctgtagcagtcagtcgtgtgacagcaatcctcaggtatcctctgtccac 780
    |||||||
DB 721 ACTACTAGTGTGACAGTGTGTGTGTGACAGCCATCTGAGGATCATCTTGTGCTTAC 780
QY 781 ccccaagggaagatcagcaagactatgtgacacagtcggtgtgttggcgaagggcgtgct 840
    |||||||
DB 781 CCCCAAGGGAAGATCAGAAACATTATGTGACAGTGGGTGTGGCAGAGGCGGTGCT 840
QY 841 gtgggtacctcgtgtcaactgtatccctctcgtgtgcttgcacatgtgtcgtgtg 900
    |||||||
DB 841 GTGGGTACCTCTGTCTACCTGATCCCTTCTCGTGGCTTGCATGAGTGTGCTGGTCTTGTG 900
QY 901 gctgggtcgtatctcgtcgtggggaagcaagctacgtgcgggggtgtttaaccgagtgctg 960
    |||||||
DB 901 GCTGGGTGATCTCTCGTGGGGAAGCCAGTAACCTGCGGGGTGTGTAAACCGAGTGTG 960
QY 961 ggaatccccaagctccatcatgtggtctacacacttcagctgtcgtgtgtctgtgagag 1020
    |||||||
DB 961 GGAATCCCCACAGCTCATGAGGCTACAACTTCAAGCTTGCAGGCTCTCTTGGAGAG 1020
QY 1021 atcatctacattgtgtcgtcgtgtgtgtatgatacgttcggagccggcaatgcatgtggc 1080
    |||||||
DB 1021 ATCATCTACATTGTGCTGTCTGTGTGTATCCCTGAGCGGCAATGGCATATTGGC 1080
QY 1081 ttccaggtccctcctcagatgtgggaactcagctgtgcacatcgtgtatgactatgct 1140
    |||||||
DB 1081 TTCCAGGTCTCTCTAGCATTTGGGGAACCTGCTTGGCATGTGATGCTCTACGCTCT 1140
QY 1141 ggtctcctgacaaggtgtgtctcctaaatcttaaaatagaaagaacatcattgagagctaaa 1200
    |||||||
DB 1141 GGTCTCCTGACAGGTTGCTCTTAATCTTAATATGGAAGAACCTCATGAGGCTTAA 1200
QY 1201 taatttgatgacaaagtttctcgtgaagttccctcattgtgtgtgtgttaa 1254
    |||||||
DB 1201 TATTTTGATGACCAAGTTTCTGGAAGTTTCTCATTTGGCTGTGATTTTAA 1254

```

RESULT 11
AF187846
LOCUS AF187846 1458 bp mRNA PRI 31-OCT-1999
DEFINITION Homo sapiens RHD type IIa protein mRNA, complete cds.
ACCESSION AF187846
VERSION AF187846.1 GI:6164858
KEYWORDS FLI-CDNA.
SOURCE human.
ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 1458)
AUTHORS Huang, C. H., Chen, Y. and Reid, M.
TITLE Human D(IIIa) erythrocytes: Rhd protein is associated with multiple
 dispersed amino acid variations
JOURNAL Am. J. Hematol. 55 (3), 139-145 (1997)
MEDLINE 97398395
PUBMED 9256293
REFERENCE 2 (bases 1 to 1458)
AUTHORS Huang, C. H.
TITLE Direct Submission
JOURNAL Submitted (18-SEP-1999) Biochemistry and Molecular Genetics, New
 York Blood Center, 310 East 67th Street, New York, NY 10021, USA
FEATURES Location/Qualifiers
 1..1458
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="1"
 /map="1p34-p36"

Query Match	Best Local Similarity	Score	DB	Length
Matches 1247	Conservative	0	Mismatches	7; Indels 0; Gaps 0;
1 atgaagctcaagttaccgagcgtctgtccggcgctgcctcgccctcttgggccctaacctg	99.1%;	1242.8;	DB 89;	Length 1458;
1 atgaagctcaagttaccgagcgtctgtccggcgctgcctcgccctcttgggccctaacctg	99.4%;	Pred. No. 0;		
1 atgaagctcaagttaccgagcgtctgtccggcgctgcctcgccctcttgggccctaacctg				
61 gaagagctcctatctcctcctctctatcttttttaccacatagagccttccttaaggat				
61 gaagagctcctatctcctcctcctctctatcttttttaccacatagagccttccttaaggat				
61 gaagagctcctatctcctcctcctcctctctatcttttttaccacatagagccttccttaaggat				
121 caaaaggagctcgtgtgacatccatcaagaattggccaagatctgacccgtatgctgagcatt				
121 caaaaggagctcgtgtgacatccatcaagaattggccaagatctgacccgtatgctgagcatt				
121 caaaaggagctcgtgtgacatccatcaagaattggccaagatctgacccgtatgctgagcatt				
181 ggccttgagcttcctacccctcagattccgagagacacagcttgagcagatgtgacctcaac				
181 ggccttgagcttcctacccctcagattccgagagacacagcttgagcagatgtgacctcaac				
181 ggccttgagcttcctacccctcagattccgagagacacagcttgagcagatgtgacctcaac				
241 cctctcaatgctgagcgttggtgtgacagttggccaatcctgtctgagcagcttccctgagcag				
241 cctctcaatgctgagcgttggtgtgacagttggccaatcctgtctgagcagcttccctgagcag				
241 cctctcaatgctgagcgttggtgtgacagttggccaatcctgtctgagcagcttccctgagcag				
301 ttcaccttcgggaagagttgatacaaacatggtcagatattccgctctgagcacaatgagtgct				
301 ttcaccttcgggaagagttgatacaaacatggtcagatattccgctctgagcacaatgagtgct				
301 ttcaccttcgggaagagttgatacaaacatggtcagatattccgctctgagcacaatgagtgct				
361 ttgtcgtgagctgatactcagttgagatgctgcttcttggggaagttcaactttgacagattggtg				
361 ttgtcgtgagctgatactcagttgagatgctgcttcttggggaagttcaactttgacagattggtg				
421 gttgagtgagctgtgtgagagttgacagcttttaaggcaacctggagagatgtgtcaatgataatc				
421 gttgagtgagctgtgtgagagttgacagcttttaaggcaacctggagagatgtgtcaatgataatc				
421 gttgagtgagctgtgtgagagttgacagcttttaaggcaacctggagagatgtgtcaatgataatc				
481 ttcaacaacagatcacacatagatacgtatgacacatcagttgttcgagccatttttgg				
481 ttcaacaacagatcacacatagatacgtatgacacatcagttgttcgagccatttttgg				
481 ttcaacaacagatcacacatagatacgtatgacacatcagttgttcgagccatttttgg				
541 ctgtctgtgagcgtgtgctgcctccaagacctctaccgagaggaacggagagataaagcaag				
541 ctgtctgtgagcgtgtgctgcctccaagacctctaccgagaggaacggagagataaagcaag				
541 ctgtctgtgagcgtgtgctgcctccaagacctctaccgagaggaacggagagataaagcaag				
601 aagagcaacagatcacacatagatacgtatgacacatcagttgttcgagccatttttgg				
601 aagagcaacagatcacacatagatacgtatgacacatcagttgttcgagccatttttgg				
601 aagagcaacagatcacacatagatacgtatgacacatcagttgttcgagccatttttgg				
661 ccaagtttcaaatctgtctcgtctgtaagaattccaatcgaagaaggaagatgacggtgttcaac				
661 ccaagtttcaaatctgtctcgtctgtaagaattccaatcgaagaaggaagatgacggtgttcaac				
661 ccaagtttcaaatctgtctcgtctgtaagaattccaatcgaagaaggaagatgacggtgttcaac				
721 acctactatgctgtgacagctgagctgtgtgtgaaagcattcctgaaggttcaactctgtgtcaac				
721 acctactatgctgtgacagctgagctgtgtgtgtgaaagcattcctgaaggttcaactctgtgtcaac				
721 acctactatgctgtgacagctgagctgtgtgtgtgaaagcattcctgaaggttcaactctgtgtcaac				

OY	781	ccccaaaggagatcagaacgaagtattgtgcaacctgtggcggtgttggccggaggagtgct	840
Db	781	CCCCAAGGAAGATCAGAAGACCTTATTGTGACAACACTGGCGGTGTTCGAGGAGCGTGCT	840
OY	841	gtcggatactcgtgtgcactgatcccctcttcocgttgtgttcgaatgtgtcgtgtctg	900
Db	841	GTTGGGTACTCGTGTGCACCTGATCCCTTCTCCGTGGCTTGGCAATGGTCTGGGTCTTG	900
OY	901	gctggagctgattccctcgttcgggggagccaagtaacctgcggagggtgtttaaccgagttcgt	960
Db	901	GCTGGGCTGATCTCCGTGGGGGAGCCAAAGTAACCTGCCGGGGGTGTGTMAACGAGTGGCTG	960
OY	961	gggattccccaaagctccatcatatgtggcgcacaacatcagaattgcgttcgagctcttggaaag	1020
Db	961	GGGATTCCCCAAGCTCCATCATGTGGGCTACACATTCACACTTGTCTGGGTCTTGGGAAG	1020
OY	1021	atcacatcaattgtctgtcgtgtgttatataccgttcgagaccgggcaaatgycatatgtgc	1080
Db	1021	ATCACATCAATTGTCTGCTGTGTGTTGATAACCGTCGAGCGGCAATGGCATGATTGGC	1080
OY	1081	ttcgaagctccctccatgaacattggggaaactcagctgtgccatcgttatgattctaagttct	1140
Db	1081	TTCGAGGCTCTCTCAGATTGGGGAACCTGACCTTGCCATCGTATGTCCTACAGCTCT	1140
OY	1141	gtctccctgcacaggttctccctaatttaaataatgaaagaccctcagagctaaa	1200
Db	1141	GGTCTCCGACAGGTTTGCTCTTAATCTCAAARATRTGGAAAAGACCCTCATGAGGCTAAA	1200
OY	1201	tattttagtaccaaagtttctctggaagttccctcattgtcgttgtgaatttaa	1254
Db	1201	TATTTGATGACCAAGTTTCTCGAAGCTTTCCTCATVTTGGCTGTGTGATTTTAA	1254
RESULT	12		
S57971	LOCUS		
DEFINITION	S57971	1251 bp mRNA	PRI 28-JUN-1993
KEYWORDS	Rh polypeptide II (clone RHP11) [human, erythroid cells, mRNA		
ACCESSION	S57971	Partial, 1251 nt.]	
VERSION	S57971.1	GI:299051	
ORGANISM	Source	human erythroid cells.	
REFERENCE		Homo sapiens	
AUTHORS		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
TITLE		Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.	
JOURNAL MEDLINE		1 (bases 1 to 1251)	
REMARK		Kajili,E., Umanishi,F., Iwamoto,S. and Ikemoto,S. Isolation of a new cDNA clone encoding an Rh polypeptide associated with the Rh blood group system Hum. Genet. 91 (2), 157-162 (1993) 93216282	
FEATURES		GenBank staff at the National Library of Medicine created this entry [NCBI glibsg 129037] from the original journal article. This sequence comes from Fig. 3. Map location: 1p34.3-36.1.	
source		location/Qualifiers	
CDS		1..1251	
		/organism="Homo sapiens"	
		/db_xref="taxon:9606"	
		1..1251	
		/partial	
		/note="Rh polypeptide II; This sequence comes from Fig. 3;"	
		authors show full coding region except for the stop codon"	
		/codon_start=1	
		/protein_id="AAB26081.1"	
		/db_xref="GI:299052"	
		/translation="MSKSYPRSVRCLPLMALTLLEALLILFFTHFDASLSDQGL	
		VASVYODDLTVMAALIGELTSFRHRSSVAFLNFMALGVOMALIDGFIQPP	
		PGKVITLFSRLATMSALISVLISADIVKVNLAOLWLVETVALGRIMVYSNI	
		FNDIYNHNMHIIYFAAYFGGITVAMCPKLPIPETEKDQTATPISLASLGALFLM	
		FWFVSNSCHLIPSPWLAWNLGVALVIAVSVYTIASSSLAHPOGKISKTYISAVLA	
		GVAFNSALCHLIPSPWLAWNLGVALVIAVSVYTIASSSLAHPOGKISKTYISAVLA	

LSLGEIIVLVLDIVGANGMIGFOVLLSIGELSLAIVIALTSGLTLGLLNKI
 WKAPHEAKYFDDVFWKFPHLAVGF"
 BASE COUNT 240 a 322 c 351 g 338 t
 ORIGIN

Query Match 98.9%; Score 1239.8; DB 97; Length 1251;
 Best Local Similarity 99.4%; Pred. No. 0;
 Matches 1244; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 atgagctctaagtaaccgcgagctgtctccgcgcgtccctctggccctaactg 60
 Db 1 ATGAGCTCTAAGTACCAGCGGTGTCTCCGGCGCTGCTGCCCCCTTGAACACTG 60
 QY 61 gaagcaactcattctcctctctatcttttaccataatgacgttctcttgaagat 120
 Db 61 GAAGCACTCTATTCTCTCTCTATTTTATCCACCATATGACGCTTCCCTTGAAGGAT 120
 QY 121 caaagaggtcgttgagcattcaatctgccaagatctgacgtgagtcgagcatt 180
 Db 121 CAAAGAGGCTGCGTGGCATCTATCAAGTTGGCCAAATCTGACCGTATGCGGCATT 180
 QY 181 ggttgagcttctcctacgtcagttccggagacacagctggagcagtggtccctaac 240
 Db 181 GGTGAGCTTCTCCCTCACCTGAGTTCCGGAGACACAGCTGAGCACTGTGGCTTCAAC 240
 QY 241 ctcttcagtctgagctgtgtgtcagtggaactcctgtctgacggttctcttgaagcag 300
 Db 241 CTCTTCATGCTGGCGCTGTGGTGTGACATGCGCAATCTGCTGGAGCGCTTCTCAATGACG 300
 QY 301 ttcctcttggaaggtgtgcatcacactgttcagatctcgtctgacacatgagct 360
 Db 301 TTCCCTCTGGGAAGGTGTGATCATCACTGTTCAGTATTTGGCGTGGCCACCATATAGTCT 360
 QY 361 ttgtcgtctgctatcagtgagatgtcttcttgaggagagtcacattggcgagtggtg 420
 Db 361 TTGTCGGTGTGTGATCTCAGCGGATGTCTTGGGGAAGGTCACTTGGCGCACTTGTG 420
 QY 421 gtatggtgtgtgtgaggtgtgacagctttaggcaactgtggagatgcatcagaatc 480
 Db 421 GTATGTTGCTGTGGAGGTGACAGCTTGTAGCACTGTGAGATGATCAGTAATATC 480
 QY 481 ttcacaacagactaaccaatgaaatgatgacatctacgtgtctgagccctattttggg 540
 Db 481 TTCAACACAGACTAACCAATGAATGATGACATCTACGTGTGCGACGCTATTTGGG 540
 QY 541 ctgtctgtgctgtgtgtcgtccgaagcctctacccgagggagagagataagatcag 600
 Db 541 CTGACTGTGCTGTGCTGCTGCCAAAGCTCTACCCGAGGAGACGAGATTAAGATCAG 600
 QY 601 acaagcaagataccagttgtctgcatctgtgggcccctctctgtgagatgtctg 660
 Db 601 ACAGCAACGATACCAATTTGTCTGCAATGCTGGCCCTCTCTTGTGTGATGTTCTGG 660
 QY 661 ccaagtttcaactctgtctgtgtagaagttccaaatcgaaagaaatgcccgtgttcaac 720
 Db 661 CCAAGTTTCAACTCTGCTGTGCTGAGAAATGCAATGCAAGAGAAATGCCGTGTCAAC 720
 QY 721 actactatgtctagcagtcagcgtgtgtgacagcaatctcaagggtatccctgtgtc 780
 Db 721 ACTACTATGTCTAGCAGTCAGCAGTGTGTGACACCAATCTCAGAGGTATCTCTTGGCTCAC 780
 QY 781 ccccaagagagatcagcaagactatgtgcaagtgccgtgtgtgagagagagcgtgtgct 840
 Db 781 CCCCAGAGAGATCAGCAAGACTTATGTGCAAGTCCGAGTGTGTGAGAGAGCGCTGCT 840
 QY 841 gtgggtacctcgtgtgtcactgcatccctctccggtgtgtccatgagtcgtgtgtgtg 900
 Db 841 GTGGGTACCTCGTGTACCTGATCCCTTCCGCTGTCGCAATGATGCTGTGTGTG 900
 QY 901 gctgggtgagatctcgtctcgggggggagcaagtaactgtccgggggtgttgaacagagtcgtg 960
 Db 901 GCTGGGTGATCTCGTCTCGGGGGAGCCAAATACCTGCGGCTGTGTGTGAACGAGTGTGCTG 960

QY 961 ggaattccccagctccatcatatgagcttacaacttcagcttgcgtggctgtcttggagag 1020
 Db 961 GGAATTCCACACAGCTCCATCATGGCTACAACTTCACGCTTGTGGGTGTCTTGGAGAG 1020
 QY 1021 atcatctacattgtcgtctgtgtcttgatcacgctcgcagcgcaatgagcatgtgac 1080
 Db 1021 ATCATCTACATTGTGCTGTGCTGTGATACCTCGAGCGGCATGATGATGATGGC 1080
 QY 1081 ttccaggtcctctcagcattggggaactcagcttggccatgtgatacgtctcagctct 1140
 Db 1081 TTCAGGTCTCTCCACCATTTGGGGAACCTGAGTTGGCATGATGATGATGATGATGATG 1140
 QY 1141 ggtctcctgagaggtgtgtctcctaactcctaataatgaggaagacatcagagctaaa 1200
 Db 1141 GGTCTCTGACAGAGTTGCTCTCTTAATCTTAATATGGAAGCACTCATAGAGGCTAAA 1200
 QY 1201 tatttgatgaccaagtttctcgaagtctcctcattgtgctgtgatt 1251
 Db 1201 TATTTGATGACCAAGTTTCTGGAAGTTTCTCTCATTTGGCTGTGATTT 1251

RESULT 13
 AB046420 1254 bp mRNA PRI 23-JAN-2001
 LOCUS
 DEFINITION Homo sapiens RHD mRNA for Rh blood group D antigen, complete cds.
 ACCESSION AB046420
 VERSION AB046420.1 GI:12381901
 KEYWORDS
 SOURCE Homo sapiens cDNA to mRNA.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
 AUTHORS Toshinori, O., Junko, T., Yoshiniko, T. and Eiiji, K.
 TITLE Homo sapiens RHD1 mRNA for Rh blood group antigen Rhd, complete cds
 JOURNAL Published Only in Database (2001) In press

REFERENCE
 AUTHORS Toshinori, O., Junko, T., Yoshiniko, T. and Eiiji, K.
 TITLE Submitted (21-JUL-2000) Omi Toshinori, Jichi Medical School, Dept.
 of Legal Medicine and Human Genetics; 3311-1 Yakushiji
 Minamikawachi, Kawachi, Tochigi 329-0498, Japan
 (E-mail:t-omi@jichi.ac.jp, URL:www.jichi.ac.jp, Tel:81-285-58-7342,
 Fax:81-285-44-4902)

FEATURES
 source Location/Qualifiers

1..1254
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 gene
 1..1254
 /gene="RHD"
 CDS
 1..1254
 /gene="RHD"

/codon_start=1
 /product="Rh blood group D antigen"
 /protein_id="BAB21261.1"
 /db_xref="GI:12381902"
 /translation="MSSKYPKRSVRLCLPLMALTLLEALILFFFTHYDSLDDKGL
 VASYQVGDITVMAAIGLFTSSFRHSWSVAFLFMALGVQAAIILDFLSQFP
 SKGVVITLFSIRLATMSALSVLISDAVLKVMADLVVAVLEVALGRVVISNI
 FNTDYNHNMHHITVFAAYRGLSVAMCLPRLEPGTDDKDOTATIPISAMLGALFLIM
 FMSVNSPLRSPIDOKRNAMPNTYALAVSVYRAIGSSSLAHQKRTISMTYSAVLA
 GGVAVGTSCHLIPSPMLAVYGLVAGLISVGAKYLPGCCNRYLGIPIHSTIMQYNSL
 LGLLGEIIVLVLDIVGANGMIGFOVLLSIGELSLAIVIALTSGLTLGLLNKI
 WKAPHEAKYFDDVFWKFPHLAVGF"
 BASE COUNT 242 a 323 c 347 g 342 t
 ORIGIN

Query Match 98.9%; Score 1239.6; DB 85; Length 1254;
 Best Local Similarity 99.3%; Pred. No. 0;
 Matches 1245; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

OY	1	atgagcctctaaagtaacccgagctgctgcgcggagctgctgcgcctctggagccctaaacatg	60
Db	1	ATGAGCCTCTAAAGTACCCGCGGCTGTGC CGGCGCTGCTCCTCTGGGCGCTTAACACTG	60
OY	61	gaagcagctctcaatctctctctctctatctttttaccacataagcgtctccttaagagat	120
Db	61	GAAGCAGCTCTCAATCTCTCTCTCTATTTTAAACCATTAAGACGCTTCCTTAAGAGAT	120
OY	121	caaaaggggctcgtggaatacctcttcaagtttggccaagaatctgaacctgatatgcygcact	180
Db	121	CAAAAGGGGCTCGTGGGATCTCTAAGTTGGCAAGATCTAAGCTGATGGCGGCCATT	180
OY	181	gacctggagctctccacactccgagttctccgagagacacagcttgagcaagtgtgacctcaac	240
Db	181	GCGTTGGGCTTCCTCACTCGAGTTTCGGAGACACAGCTGGACGAGTGTGGCCTTCAAC	240
OY	241	ctcttcatactgctgcgcttggatgtgcagttgggcaatccctgcctggaacgcttccctgaagcag	300
Db	241	CTCTTCATAGCTGGGGCTTGGTGTGCAGTGGGCAATCTCTGAGGCGCTTCCTGAAGCAG	300
OY	301	ttccctctctggaaagtggtgcatacacacgtttcaagttacgttcgctgccaacataagttact	360
Db	301	TTCCCTCTCTGGGAAGTGGGTCAATCACAGTTCAGTATTCGGCTGGCCACCATGAGTCT	360
OY	361	cttgctgctgctgatactcagtagtgaatgctgcttctgggaaagttcaacttggcgcagttggtg	420
Db	361	CTTGCTGCTGCTGATCTCAGTGAATGCTGCTTGGGAAAGTCAACTGGCGCAGTTGGTGTG	420
OY	421	gttagttgtctcgtgtggaggtggaacagtttaggaacacttagatgtgataagtaataatc	480
Db	421	GTTAGTTGTCTCGTGTGGAGGTGAACAGTTTAGGAACCTTAGATGTGATACGTAATATC	480
OY	481	ttcaacaacagactacccacatagaaacatgatatgacatctagctgtctgcagactattttggg	540
Db	481	TTCAACAACAGACTCCACATGAACATGAATGACATCTAGTGTGCGACGCTTAATTTGGG	540
OY	541	ctgtgctgtgctcgtgtgctgcgccaagcctctaccccgaaaggaaaggagataaagaatcag	600
Db	541	CTGTGCTGTGCTGTGGTCTGTGCCAAAGCCTCTACCCGAAGGAACGAGAGATTAAGATAG	600
OY	601	acaagaacagatacccaattgtctctgcacatgctgagcgccctctctcttggatgtgtctcgg	660
Db	601	ACAGACAAGATACCCAACTTTGTGTGCCAATGCTGTGGGCGCCTCTCTGTGGATGTTTCTGG	660
OY	661	ccaagttctcaactctgctctgtctgagaagtccaatcgaaaggaagaatgycgtgtgtcaac	720
Db	661	CCAAGTGTCAAACTCTCTCTGTCTGTGAAGATCCAAATCCAAAGGAAGAATGCCATGTTCAAC	720
OY	721	accacacatgctgtagaagtcagcgtgtgtgaaacagccaatacgaagtgtaactctgtgctaac	780
Db	721	ACCTACATGTCTTAGCAGATCAAGTGTGTGACAGCCATCTCAGGGTCAATCCTTGACTAC	780
OY	781	cccccaaggagaatcagcaagaacttattgtgcacagtgcgtgtgttgccaggaagcgttgct	840
Db	781	CCCCAAAGGAAGATCACACATGACTTATGTGTGACAGTGCAGTGTGGCAGGAAGCGTGGCT	840
OY	841	gttggtgtaacctcgtgtcaacactgacccttctccgttggtctgtcgaatggtgtctggtctctg	900
Db	841	GTGGTGTAACTCTGTCAKCCGTGATCCCTTCCGTGGCTTCCATGATGTGTGTGGTCTTGTG	900
OY	901	gcttgagctgatactcgtctcggaggagacaagtaacctgcggagtgctgtgtaaccgaatgtgtg	960
Db	901	GCTTGAGCTGATCTCCGTCTGGGGAGGCAAGTACCTCCGGGGTGTGTAAACGAGATGCTG	960
OY	961	ggagatccccaacagctccacatcatgtggcctacaactcaagctgtgcctggtctgctgtggagag	1020
Db	961	GGGATTCCTCCACAGCTCCATCATGAGGCTCAACAATTCACTGCTGGGCTGCTGTTGAGAG	1020
OY	1021	atacatcaaatgtgctgctggtgtgtctgtaatacgtctggagccgggaaatgagcatgattggc	1080
Db	1021	ATCATCTACATATGTGTGCTGTGGCTTGTGAATACGCTGTGGAGCGGGAATGGCATGATTGGC	1080

[illegible]

QY	121	caaaaggagccgfcgagcaaccatcaagttggccaagaatctggaccgfyagagcgccatt	180
Db	121	CAAAAGGAGCCGFCGAGCAACCTTANCAAGTGGCCAAABATTCGACCGTGAATGGGCCCTT	180
QY	181	gacctgggcttcctccacctcgaagtttccggagagacacagctggagcagttgtgacctcaac	240
Db	181	GGCTTGGGCTTCCTCACCTCAAAATTTCCGGAGACAGCTGGAGCAATGTGGCTTCAC	240
QY	241	ctcttcacatgctgcgcttgbtgbtgcagtyggacaatccctgctgagcggctctcagaacag	300
Db	241	CTCTTCATATGCTGGCGCTTGGGTGTCAGATGGGCAATCTCGCTGGAGCGCTTCAGACAG	300
QY	301	ttccctctcgggaaggtggtatcaacactgttcagttatcttgctgtgcccacatgaatgct	360
Db	301	TTTTCCCTCCTGGGAAGGTGGTATCACACTGTTCAGTATTTGGCGGGCCACATGAGTGTCT	360
QY	361	ttgcgcgtygcgcgtactgaatgagtcgctctcttgggaaggttcaacttggcgacagttggtg	420
Db	361	ATGTCGGTGGCTGAATCTACGCGGGTGTCTTGGGGAAAGTCAACTTGGCGCAAGTTGGTG	420
QY	421	gtgatygtcgtgtygagagtygacagctttaggcaaccttggagatgtygtcatcaatatac	480
Db	421	GTGATGGTGGCTGGTGGAGGTGACAGCTTTTAAAGCACCTTGAGGATGGTCATCAATATC	480
QY	481	ttcaacacagcactcacatgaaatgatatgacatactgaagtgtgtcgaacctattttggg	540
Db	481	TTTCACACAGACTCCACATGAAATGATATGACATCTACGTATGGCTGGCGACCTATTTTGGG	540
QY	541	ctgctcgtgagcctgtgagccgtccaaagccctctaccgcgaaggaaagagagataaagatcag	600
Db	541	CTGCTCTGAGCCTGTGTGCTGCCAACACCTCTTACCCGAGGAAACGAGATTAAGATCAG	600
QY	601	acagacaagatacccaagtttgbtctgcgaatgctggcgccctctctgtgtgaatgtctgg	660
Db	601	ACAGACAAGATACCACAGTTTGTGTGCATGGTGGGCCCCCTCTCTGTGTGATGTTCTGG	660
QY	661	ccaaagtttcaactctgtctgtcgtgtagaagaatcccaatggaagaaagaaatgcgcgtttcaac	720
Db	661	CCAAAGTTTCAACTCTGTCTGTCTGTAGAGATGCCAAATTCAAAAGAAAGATGCCGTGTTTCAC	720
QY	721	acctactatgtctagcagatcaagtggtgtgtacacacatcccaaggtatcccttggccac	780
Db	721	ACCTACTATGTGCTGTAACGATGAGCGTGTGCACACCAATCTCAAGGCTATCCTTGGCTCAC	780
QY	781	ccccaaagggaagatcaagacaagacttatgtgcacagtygcagtygttctgcaggaagcgtygct	840
Db	781	CCCCAAAGGAAGATCAGACAACACTTATGTGCACAGTGGGTTGGCAGGAGCGGTGCT	840
QY	841	gtggagtacactgtgtacactgtatccctctccctgctgtgcttgcacatgtgtcgtggtgtg	900
Db	841	GTGGAGTACCTGTGTACACTGTATCCCTTCCCGTGGCTTCCATGTGCTGGGCTTGTGTG	900
QY	901	gctggagctgacatctcgcgtccggggggccaagttacctgcgggggtygtgttaacgagtygtg	960
Db	901	GCTGGAGCTGAATCTCGTGGGGGACCAAGTCTGCGGGGGTGTTTAAACCAAGTGTCTG	960
QY	961	ggagattcccccacagctccatcatatgtgtacaaacttcaagcttgcctgggtctctgtcttggag	1020
Db	961	GGGATTCGCCACACTCCATCATGTGGGTACAACTTCAGTTGCTGGGTGCTGTTGGAGAG	1020
QY	1021	atcatctacatltgtcgtcgtggtgtctgtatacogtgcggagccggcaatgtgcatgtatgctc	1080
Db	1021	ATCATCTACATLTGTGTGTGTGCTGTGATACCTCTCGAGACGGCAATGTGCATGATTTGGC	1080
QY	1081	ttcagaagtcctccctcagaatctgggaactgaagcttggcacatcgtgatagcttccaagttc	1140
Db	1081	TTCCAGTCTCTCCTCAGATTTGGGAACTCAGCTTGGCCATCGGATGAGCTCTCACGTCT	1140
QY	1141	ggtcctccygaagaagtttgcctcctaacttaanaatagyaagaagaccctcatatgtgcctaaa	1200
Db	1141	GGTCTCCCGACAGGTTTGCTCTTAATCTTAAATATAGGAAGCACTCATAGAGGCTTAAA	1200
QY	1201	tatttgcatagcacaagtttctctggaagtttcccatattgctcgttgtatatttaa	1254

DB	1201	TATTTTGATGACCAAGTTTCTTCGAGAGTTTCTCTCATTTGGCTGTGGATTTTAA	1254
RESULT	15		
LOCUS	AF037626	1247 bp	mRNA
DEFINITION	Homo sapiens truncated RHD (RHD mutant) mRNA, complete cds.		
ACCESSION	AF037626		
VERSION	AF037626.1	GI:4104640	
KEYWORDS			
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
AUTHORS	1 (bases 1 to 1247)		
TITLE	Andrews, K.A., Walter, L.C., Saul, A. and Hyland, C.A.		
JOURNAL	The Rh D antigen negative trait in an Rh CDee phenotypic Caucasian attributed to a four nucleotide deletion in the Rh D gene		
REFERENCE	Unpublished		
AUTHORS	2 (bases 1 to 1247)		
TITLE	Andrews, K.A. and Hyland, C.A.		
JOURNAL	Submitted		
REFERENCE	Submitted (10-DEC-1997) Malaria and Arbovirus Unit, Queensland Institute of Medical Research, 300 Herston road, Herston, QLD 4006, Australia		
FEATURES			
source	Location/Qualifiers		
	1..1247		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/chromosome="1"		
	/map="1p34.3-p36.1"		
	1..1247		
gene	/gene="RHD"		
	/note="RHD gene with a four nucleotide deletion at the state of exon 4, resulting in a D antigen negative phenotype"		
	1..498		
CDS	/gene="RHD"		
	/codon_start=1		
	/product="truncated Rhd"		
	/protein_id="AAD02100.1"		
	/db_xref="GI:4104641"		
	/translation="MSSKRPKSVRCRLPLMALTEAALLLEFFTHYDASLEDDKGLVASVYQODILVMAAIGLIGLFTSFRRHSMSSVAFEMIALGVQMLIDGFLSOPPSGKVIITLFSIRLATMALSVLISVDVAQLKVNLAQLVYVIVEVTALGNLRWISNINFTNT"		
BASE COUNT	237 a 322 c 349 g 339 t		
ORIGIN			
Query Match	98.2% Score 1231.4; DB 88; Length 1247;		
Best Local Similarity	99.6% Pred. No. 0; Mismatches 1; Indels 4; Gaps 1;		
Matches 1246; Conservative	0;		
QY	1 atgagctctaaagctaccgacgagctctccgagcgctcctgcgcccctctggccctaacactg 60		
DB	1 ATGAGCTCTAAAGTACCACCGAGCTCTCCGAGGCGTCCCTGCCCCCTTGGCCCTAACA CTG 60		
QY	61 gaagcagctctcaattctccctctctattttttttaccacatagacgttccttagagat 120		
DB	61 GAAGCAGCTCTCATATCTCTCTCTCTATATTTTATTTTACCACCACTATGAGCGTTCCTTAAGAGAT 120		
QY	121 caaagggagctcgtagcagcctcatcaatgaattggcagaagtctgacgctatagcgagca 180		
DB	121 CAAAAGGGGAGCTCGTAGCACTCATCAAGCTTGAGCAAGATCTGACCGTATAGGCGCAATT 180		
QY	181 ggcctggagctctctcaactcagattccggagacacaagctggagcagtggtgacctcaac 240		
DB	181 GGCTTGGGCTTCCCTCACTCGAGTTCGCGGAGACACACTGGAGCAGTAGTGCCCTTCAAC 240		
QY	241 ccttcacgtctggcgcttggtgctgagctggagcaatcctgctgagacgctctccctgagcag 300		

Db 241 CTCTTCAATGCTGGCGCTTGCTGTGCAGTGGCAATCCTGCTGACAGGCTTCTGAGCCAG 300
QY 301 ttccctctctggaaggtgcatcacatgttcaagtaatcgtctgagccacatgagtgtc 360
|||||
Db 301 TTCCCTTCTGGGAAGGTGATCAGCTGTTCAGTATTCGGCTGGCCACATGAGTGTCT 360
QY 361 ttgtcggtctgatactcagtgatgtctgtctggggaagtcacactggcgagttgtg 420
|||||
Db 361 TTGTGGGTGCTGATCTCAGTGGATGCTGTGGGGAAGTCAACTTGGCGCAGTTGGTG 420
QY 421 gtatggtgtgtgtggaagtgagacagctttaggaaccttagagagatggtcatcaatatac 480
|||||
Db 421 GTATGAGTGTGCTGGTGGAGTGACAGCTTTAGGCACTGAGGATGATCATCACTAATATC 480
QY 481 ttcaacacagactacacacatgaaatgatacatcagtgatcagtgatcagcctaatttggg 540
|||||
Db 481 TTCAAC----ACTACCAATGAAATGATGATGATGATGATGATGATGATGATGATGATGATG 536
QY 541 ctgtctgtggtctgt 600
|||||
Db 537 CTGTCTGTGTGCTGT 536
QY 601 aaggaacagatcagcagcttctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 660
|||||
Db 597 ACAGCAACGATACCACTTGT 656
QY 661 ccaagttcacaactctgtctgt 720
|||||
Db 657 CCAAGTTTCAACTGT 716
QY 721 acctactatgtctgt 780
|||||
Db 717 ACCTACTATGCTGT 776
QY 781 ccccaaggaagatacgaacagacttattgtacagtgagtgatgtgtgtgtgtgtgtgtgt 840
|||||
Db 777 CCCCAAGGAAGATCAGCAAGACTTATGTCAAGTGTGTGTGTGTGTGTGTGTGTGTGTGT 836
QY 841 gtgtgtacctgt 900
|||||
Db 837 GTGGGTACCTGT 896
QY 901 gctgt 960
|||||
Db 897 GCTGT 956
QY 961 ggaattcccaacagctccacatcagtgagtgagtgagtgagtgagtgagtgagtgag 1020
|||||
Db 957 GGGATTCGCCACAGCTGCATCATGAGGCTACAACTTCACTTGTGTGTGTGTGTGTGTGT 1016
QY 1021 atcatctacatgt 1080
|||||
Db 1017 ATCATCTACATGT 1076
QY 1081 ttccaggtctctctcagcatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1140
|||||
Db 1077 TTCAGGTCTCTCAGCATTTGGGAATCAGCTTGGCCATCTGTGTGTGTGTGTGTGTGT 1136
QY 1141 ggtctcctgacaggt 1200
|||||
Db 1137 GGTCTCTGACAGGT 1196
QY 1201 taatttgatgacaaagtttctctgaagtttctcatttgagttgtgtgtgtgtgtgt 1251
|||||
Db 1197 TATTTTGATGACCAAGTTTCTGGAAGTTTCTCATTTGTGCTGTGTGTGTGTGTGTGT 1247

!!AA_SEQUENCE 1.0
 ID AAY24056 standard; Protein; 417 AA.
 AC AAY24056;
 DT 04-OCT-1999 (first entry)

DE Protein encoded by the prevalent allele of the Rhd gene.

XX Allele; Rhesus D antigen; Rhd; weak D phenotype; blood transfusion.

OS Homo sapiens.

PN W09937763-A2.

PD 29-JUL-1999.

PF 18-DEC-1998; 98WO-EP08319.

PR 23-JAN-1998; 98EP-0101203.

PA (DRKB-) DRK BLUTSPENDEDIENST BADEN WÜRTTEMBERG.

PI Flegel WA, Wagner FF;

DR WPI; 1999-469127/39.

DR N-PSDB; AAX86522.

PT Nucleic acid sequences correlated with Rhesus weak D phenotype,
 useful for screening blood from donors and recipients for
 transfusion methods

PS Disclosure; Fig 2: 64pp; English.

CC The present sequence is encoded by the prevalent allele of the Rhesus D
 CC (Rhd) antigen gene. The specification describes a Rhd contributing to
 CC or indicative of the weak D phenotype, where the Rhd polynucleotide
 CC carries at least one missense mutation as compared to the wild-type Rhd,
 CC in its transmembrane and/or intracellular regions, especially in amino
 CC acid positions 2-16, 114-149, 179-225 or/and 267-397, with the proviso
 CC that the D antigen does not carry a single missense mutation leading to
 CC a F223V or T283I substitution. The probes and antibodies are useful in
 CC the methods for detection of weak D phenotypes. Red blood cells, from
 CC probands, are useful for the assessment of the affinity, avidity and/or
 CC reactivity of monoclonal anti-D antibodies, polyclonal anti-D antisera or
 CC of anti-globulin or anti-human globulin antisera. Detecting the presence
 CC of the Rhd associated with weak D phenotype is useful for determining
 CC that a patient in need of a blood transfusion is to be transfused with
 CC Rhd negative blood from a donor. Alternatively, testing for weak D
 CC phenotype Rhd in the blood of a donor is useful for determining whether
 CC the donor blood should be excluded for transfusion to patients having
 CC wild type Rhd or weak D types, other than that of the donor weak D
 CC type.

CC Sequence 417 AA;
 SQ

AAZ24056 Length: 451 September 11, 2001 09:20 Type: P Check: 4689

1 SQARNDBCQE ZGHSQILKMF PSTWYVSQOT HERSSMSKYP RSVRCLPLM
 51 ALTLEALIL LFYFTHYDA SLEDOKGLVA SYVGQDLTV MAALGLGLF
 101 SSFRRHSWSS VAFNLFMAL GYQNALLDG FLSPQPSGKV VTLFSIRLA
 151 TMSALSVLIS VDAVLGKVNl AOLVVMVLVE VTALGNLRMV ISNIFNTDYH
 201 MNMHIYVFA AYFGLSVAMC LPKPLPEGTE DKDQTATIPS LSAMLGALFL
 251 WMFPSPNSA LRSPIERKN AVNTIYAVA VSVYTAISGS SLAHPCKIS
 301 KTYVHSAVLA GGYAVGTSCH LIPSPWLMV LGLVAGLISV GGAKEYLPGCC
 351 NRVLGIPHSS IMGYNFSLG ILGEIITYVL LVLDTVGAGN GMIGFVLLS

401 IGEISLAIVL ALTSGILTGL LNLKIMKAP HEAKYEDQV FKKFPLAVG
 451 F

THIS PAGE BLANK (USPTO)

```
; Sequence 41, Application US/09600714
; GENERAL INFORMATION:
; APPLICANT: Flegel, Willy A.
; APPLICANT: Wagner, Franz F.
; APPLICANT: DRK Blutspendedienst Baden-Wuerttemberg gGmbH
; TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES CORRELATED
; TITLE OF INVENTION: WITH THE RHESUS WEAK D PHENOTYPE
; FILE REFERENCE: 12086-002001
; CURRENT APPLICATION NUMBER: US/09/600,714
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: PCT/EP98/08319
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 41
; LENGTH: 1254
; TYPE: DNA
; ORGANISM: Homosapien
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1251)
09-600714-1
MSKRYRSVRRCPLPMALEALILLFYFFTHYDASLEDQKGLVASYOGDLYTMAAIGLGLTSR
RHSWSSVAENLEMLALGVQWAILLDGFLSQFPGKAVITLFSIRLATMSALSVLISVDVGLGVNLAQV
VMVLVEVTALGNLRNVIISNIFNTDYMNMNMHIYFAAYFGLAYMCLPKPLPEGTEDKQATATIPSLSAM
LGLALFLMMPSPNSALRSPIERKNVAVNTYAVAVSVYTAISGSLAHPOGKISKTYGHSVILPBGVA
VGTSCHLIPSPMLAVIAGLIVAGLISVGAKEYLPGCCNRVILGPHSSIMGYNFSLGLLEIYIVLVLD
TVGAGNGMIGFOVLSIGELSLAIVIALTSGLLTGLLNLKIKAPHEAKYFPDDQYFWKFPHLAVGFXI
```

```
; Sequence 41, Application US/09600714
; GENERAL INFORMATION:
; APPLICANT: Flegel, Willy A.
; APPLICANT: Wagner, Franz F.
; APPLICANT: DRK Blutspendedienst Baden-Wuerttemberg gGmbH
; TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES CORRELATED
; TITLE OF INVENTION: WITH THE RHESUS WEAK D PHENOTYPE
; FILE REFERENCE: 12086-002001
; CURRENT APPLICATION NUMBER: US/09/600,714
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: PCT/EP98/08319
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 41
; LENGTH: 1254
; TYPE: DNA
; ORGANISM: Homosapien
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1251)
09-600714-2
MSKRYRSVRRCPLPMALEALILLFYFFTHYDASLEDQKGLVASYOGDLYTMAAIGLGLTSR
RHSWSSVAENLEMLALGVQWAILLDGFLSQFPGKAVITLFSIRLATMSALSVLISVDVGLGVNLAQV
VMVLVEVTALGNLRNVIISNIFNTDYMNMNMHIYFAAYFGLAYMCLPKPLPEGTEDKQATATIPSLSAM
LGLALFLMMPSPNSALRSPIERKNVAVNTYAVAVSVYTAISGSLAHPOGKISKTYGHSVILPBGVA
VGTSCHLIPSPMLAVIAGLIVAGLISVGAKEYLPGCCNRVILGPHSSIMGYNFSLGLLEIYIVLVLD
TVGAGNGMIGFOVLSIGELSLAIVIALTSGLLTGLLNLKIKAPHEAKYFPDDQYFWKFPHLAVGFXI
```

THIS PAGE BLANK (USPTO)

RESULT 2

A: Molecule type: protein Rhd - human
A: Residues: 401-407, 'DI' <SUY>
R: Westhoff, C.M.; Wylie, D.E.
Blood 83, 3098-3100, 1994
A: Title: Identification of a new Rhd-specific mRNA from K562 cells.
A: Reference number: 152589; PMID:94235883
A: Accession: 152589

A: Status: preliminary; translated from GB/EMBL/DBJ
A: Molecule type: mRNA
A: Residues: 1-217, 'M', 219-417 <RE2>
A: Cross-references: GB:S70174; NID:9546725; PIDN:AAB30756.1; PID:9546726
R: Suyama, K.; Lunn, R.; Haller, S.; Goldstein, J.
Blood 84, 1975-1981, 1994
A: Title: Rh(D) antigen expression and isolation of a new Rh(D) cDNA isoform in human
A: Reference number: 152596; PMID:94362249
A: Accession: 152596

A: Status: preliminary; translated from GB/EMBL/DBJ
A: Molecule type: mRNA
A: Residues: 1-217, 'M', 219-417 <RE3>
A: Cross-references: GB:S73913; NID:9693866; PIDN:AAB31911.1; PID:9693867
C: Genetics:
A: Gene: GDB:RHD; RH; RHD
A: Cross-references: GDB:119551; OMIM:111680
A: Map position: 1p36.2-1p34
C: Keywords: erythrocyte; surface antigen; transmembrane protein

A: Status: preliminary; nucleic acid sequence not shown; translation not shown
A: Accession: 526564
A: Reference number: 526560; PMID:92360855

A: Status: preliminary; nucleic acid sequence not shown; translation not shown
A: Molecule type: mRNA
A: Residues: 1-417 <LEV2>
A: Cross-references: EMBL:X63097; GB:S48953; NID:936046; PIDN:CAA44811.1; PID:936047
A: Experimental source: bone marrow, clone RHXIII
A: Note: sequence extracted from NCBI backbone (NCBI:118725, NCBI:P.118726)
R: Le Van Kim, C.; Cherif-Zahar, B.; Raynal, V.; Mouro, I.; Lopez, M.; Cartton, J.P.; Col
Blood 80, 1074-1078, 1992
A: Title: Multiple Rh messenger RNA isoforms are produced by alternative splicing.
A: Reference number: 526560; PMID:92360855
A: Accession: 526564

A: Status: preliminary; nucleic acid sequence not shown; translation not shown
A: Molecule type: mRNA
A: Residues: 1-217, 'M', 219-417 <LEV3>
A: Cross-references: EMBL:X63097; NID:936027; PIDN:CAA44808.1; PID:936028
A: Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1991
R: Ace, M.A.; Thompson, E.S.; Wagner, S.; Coyne, K.E.; Ferdman, B.A.; Lublin, D.M.
Blood 82, 651-655, 1993
A: Title: Molecular cloning of Rhd cDNA derived from a gene present in Rhd-positive, but
A: Reference number: 152581; PMID:93320449
A: Accession: 152581

A: Status: preliminary; translated from GB/EMBL/DBJ
A: Accession: 502085
A: Molecule type: protein
A: Residues: 1-217, 'M', 219-417 <RES>
A: Cross-references: GB:I08429; NID:9337390; PIDN:AAA02679.1; PID:9337391
R: Arent, N.D.; Ridgwell, K.; Newby, W.J.; Tanner, M.J.A.; Anstee, D.J.; Kumpel, B.
Biochem. J. 256, 1043-1046, 1988
A: Title: Protein-sequence studies on Rh-related polypeptides suggest the presence of at
A: Reference number: 502085; PMID:89134163
A: Accession: 502085

A: Molecule type: protein
A: Residues: 2-11, 'X', 13-33 <AVE1>
A: Note: this peptide was designated blood group Rh-related protein D30
A: Accession: 502087

A: Molecule type: protein
A: Residues: 2-11, 'X', 13-14 <AVE2>
A: Note: this peptide was designated blood group Rh-related protein R6A32
R: Bloy, C.; Blanchard, D.; Dahn, W.; Beyreuther, K.; Salmon, C.; Cartton, J.P.
Blood 72, 661-666, 1988
A: Title: Determination of the N-terminal sequence of human red cell Rh(D) polypeptide an
A: Reference number: A32509; PMID:88294325
A: Accession: A32509

A: Molecule type: protein
A: Residues: 2-11, 'L', 13-15, 'X', 17 <BL0>
R: Saboori, A.M.; Smith, B.L.; Agre, P.
Proc. Natl. Acad. Sci. U.S.A. 85, 4042-4045, 1988
A: Title: Polymorphism in the M-r 32,000 Rh protein purified from Rh(D)-positive and -neg
A: Reference number: A30216; PMID:88234555
A: Accession: A30216

A: Molecule type: protein
A: Residues: 2-11, 'X', 13-21 <SAB>
R: Suyama, K.; Goldstein, J.; Abersold, R.; Kent, S.
Blood 77, 411, 1991
A: Title: Regarding the size of Rh proteins.
A: Reference number: A61174; PMID:91091537
A: Accession: A61174

A: Molecule type: protein
A: Residues: 401-407, 'DI' <SUY>
R: Westhoff, C.M.; Wylie, D.E.
Blood 83, 3098-3100, 1994
A: Title: Identification of a new Rhd-specific mRNA from K562 cells.
A: Reference number: 152589; PMID:94235883
A: Accession: 152589

A: Status: preliminary; translated from GB/EMBL/DBJ
A: Molecule type: mRNA
A: Residues: 1-217, 'M', 219-417 <RE2>
A: Cross-references: GB:S70174; NID:9546725; PIDN:AAB30756.1; PID:9546726
R: Suyama, K.; Lunn, R.; Haller, S.; Goldstein, J.
Blood 84, 1975-1981, 1994
A: Title: Rh(D) antigen expression and isolation of a new Rh(D) cDNA isoform in human
A: Reference number: 152596; PMID:94362249
A: Accession: 152596

A: Status: preliminary; translated from GB/EMBL/DBJ
A: Molecule type: mRNA
A: Residues: 1-217, 'M', 219-417 <RE3>
A: Cross-references: GB:S73913; NID:9693866; PIDN:AAB31911.1; PID:9693867
C: Genetics:
A: Gene: GDB:RHD; RH; RHD
A: Cross-references: GDB:119551; OMIM:111680
A: Map position: 1p36.2-1p34
C: Keywords: erythrocyte; surface antigen; transmembrane protein

Query Match 93.0%; Score 1976; DB 2; Length 417;
Best Local Similarity 95.0%; Pred. No. 1,7e-152;
Matches 396; Conservative 4; Mismatches 17; Indels 0; Gaps 0;

QY 1 MSCKPRSVGRCLPLCALTEALILLEFFPTHYDASLEDOKGLVASYOGDLYMAAT 60
DB 1 MSCKPRSVGRCLPLMALTEALILLEFFPTHYDASLEDOKGLVASYOGDLYMAAT 60

QY 61 GLGFLTSSFRHSMSSVAFLMFLALGVOMAILLDGFLSQFSGKVVITLFSIRLTMSA 120
DB 61 GLGFLTSSFRHSMSSVAFLMFLALGVOMAILLDGFLSQFSGKVVITLFSIRLTMSA 120

QY 121 LSVLISVDALGVKNLAQVVMVLEVTGLGNRMVSNFTNDYHNMNHIVFAAYEG 180
DB 121 LSVLISVDALGVKNLAQVVMVLEVTGLGNRMVSNFTNDYHNMNHIVFAAYEG 180

QY 181 LVAWCLPKPLPCTEDNDQATPISLAMLGLFIMFSPNSALLRSPERKNAVEN 240
DB 181 LVAWCLPKPLPCTEDNDQATPISLAMLGLFIMFSPNSALLRSPERKNAVEN 240

QY 241 TYRAVANSVYTAISGSSLAHPQKISTYGHSAVLPREGAVDTSCHLIPRPIVLGLV 300
DB 241 TYRAVANSVYTAISGSSLAHPQKISTYGHSAVLPREGAVDTSCHLIPRPIVLGLV 300

QY 301 AGLISVGAATYLPCCGKRVVIGPHSSIMGNFSLGLLEIYIVLLDPTVGAGNGMIG 360
DB 301 AGLISVGAATYLPCCGKRVVIGPHSSIMGNFSLGLLEIYIVLLDPTVGAGNGMIG 360

QY 361 FOYVLISIGELSLAIVALTSGLLTALLNLKIRKAPHEAKYFPDQYFWKPRHLAVG 417
DB 361 FOYVLISIGELSLAIVALTSGLLTALLNLKIRKAPHEAKYFPDQYFWKPRHLAVG 417

RESULT 3

Rhesus-like protein - chimpanzee
C: Species: Pan troglodytes (chimpanzee)
C: Date: 04-Oct-1996 #sequence_revision 04-Oct-1996 #text_change 21-Jul-2000
C: Accession: I37005

R: Salvagnol, I.; Blancher, A.; Calvas, P.; Clayton, J.; Socha, W.W.; Collin, Y.; Ruffi
Biochem. Genet. 32, 201-221, 1994
A: Title: Molecular genetics of chimpanzee Rh-related genes: their relationship with t
A: Reference number: I37003; PMID:95085595
A: Accession: I37005

A: Status: preliminary; translated from GB/EMBL/DBJ
A: Molecule type: mRNA

A:Residues: 1-417 <RES>
A:Cross-references: GB:I37050; NID:g606991; PIDN:AAA65624.1; PID:g606992

Query Match 87.7%; Score 1863; DB 2; Length 417;
Best Local Similarity 89.2%; Pred. No. 2,4e-143;
Matches 37/2; Conservative 12; Mismatches 33; Indels 0; Gaps 0;

```

1 MSCKYPRSVGRCLPLCALTEALILLFYFTHYDASLEDDKGLVASYOVGODLTVAAI 60
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1 MSCKYPRSVGRCLPLCALTEALILLFYFTHYDASLEDDKGLVASYOVGODLTVAAI 60
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 61 GLGFLTSSFRHSSWSVAFLFMLALGVOMAILDGLFLOSPPSGKVITTFISILATMSA 120
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 61 GFGFLTSSFRHSSWSVAFLFMLALGVOMAILDGLFLOSPPSGKVITTFISILATMSA 120
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 121 LSVLISDAVLGKYNLAQLVVMVLEVTDLGNLRMVISNIFNTDYHNNMMHIYFAAYFG 180
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 121 LSVLISDAVLGKYNLAQLVVMVLEVTDLGNLRMVISNIFNTDYHNNMMHIYFAAYFG 180
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 181 LTVAMCLPKPLPECTEDNDQATIPSLSAMLGALFLMFPSPVSALLRSPIERKNVFN 240
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 181 LTVAMCLPKPLPECTEDNDQATIPSLSAMLGALFLMFPSPVSALLRSPIERKNVFN 240
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 241 TYAVAVSVVTAISGSSSLAHPOGKISKTYGHSVAVLPEGVAVDTSCHLIPSPWLPVLGLV 300
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 241 TYAVAVSVVTAISGSSSLAHPOGKISKTYGHSVAVLPEGVAVDTSCHLIPSPWLPVLGLV 300
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 301 AGLISVYGAKYLPCCNRRVLGIPHSSTMGYNFSLGLLEIITYVLVLDTVGANGMIG 360
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 301 AGLISVYGAKYLPCCNRRVLGIPHSSTMGYNFSLGLLEIITYVLVLDTVGANGMIG 360
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 361 FOVLSTIGELSLAIVLTGSLTGLTALLNKIRKAPHEAKYFDQVFWKPEHLAVGF 417
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 361 FOVLSTIGELSLAIVLTGSLTGLTALLNKIRKAPHEAKYFDQVFWKPEHLAVGF 417
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

RESULT 4

Rhesus-like protein - gorilla
C:Species: Gorilla gorilla (gorilla)
C>Date: 04-Oct-1996 #sequence_revision 04-Oct-1996 #text_change 21-Jul-2000
C/Accession: I37076
R:Salviagnol, I.; Blancher, A.; Calvas, P.; Clayton, J.; Socha, W.W.; Colln, Y.; Ruffe,
Biochem. Genet. 32, 201-221, 1994
A:Title: Molecular genetics of chimpanzee Rh-related genes: their relationship with the
A:Reference number: I37003; MUID:95085395
A:Accession: I37076
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-417 <RES>
A:Cross-references: GB:I37053; NID:g607011; PIDN:AAA65627.1; PID:g607012

Query Match 86.7%; Score 1842; DB 2; Length 417;
Best Local Similarity 89.0%; Pred. No. 1.2e-141;
Matches 37/1; Conservative 11; Mismatches 35; Indels 0; Gaps 0;

```

1 MSCKYPRSVGRCLPLCALTEALILLFYFTHYDASLEDDKGLVASYOVGODLTVAAI 60
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1 MSCKYPRSVGRCLPLCALTEALILLFYFTHYDASLEDDKGLVASYOVGODLTVAAI 60
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 61 GLGFLTSSFRHSSWSVAFLFMLALGVOMAILDGLFLOSPPSGKVITTFISILATMSA 120
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 61 GFGFLTSSFRHSSWSVAFLFMLALGVOMAILDGLFLOSPPSGKVITTFISILATMSA 120
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 121 LSVLISDAVLGKYNLAQLVVMVLEVTDLGNLRMVISNIFNTDYHNNMMHIYFAAYFG 180
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 121 LSVLISDAVLGKYNLAQLVVMVLEVTDLGNLRMVISNIFNTDYHNNMMHIYFAAYFG 180
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 181 LTVAMCLPKPLPECTEDNDQATIPSLSAMLGALFLMFPSPVSALLRSPIERKNVFN 240
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 181 LTVAMCLPKPLPECTEDNDQATIPSLSAMLGALFLMFPSPVSALLRSPIERKNVFN 240
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

QY 241 TYAVAVSVVTAISGSSSLAHPOGKISKTYGHSVAVLPEGVAVDTSCHLIPSPWLPVLGLV 300
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 241 TYAVAVSVVTAISGSSSLAHPOGKISKTYGHSVAVLPEGVAVDTSCHLIPSPWLPVLGLV 300
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 301 AGLISVYGAKYLPCCNRRVLGIPHSSTMGYNFSLGLLEIITYVLVLDTVGANGMIG 360
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 301 AGLISVYGAKYLPCCNRRVLGIPHSSTMGYNFSLGLLEIITYVLVLDTVGANGMIG 360
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 361 FOVLSTIGELSLAIVLTGSLTGLTALLNKIRKAPHEAKYFDQVFWKPEHLAVGF 417
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 361 FOVLSTIGELSLAIVLTGSLTGLTALLNKIRKAPHEAKYFDQVFWKPEHLAVGF 417
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

RESULT 5

Rhesus-like protein - chimpanzee
C:Species: Pan troglodytes (chimpanzee)
C>Date: 04-Oct-1996 #sequence_revision 04-Oct-1996 #text_change 21-Jul-2000
C/Accession: I37004
R:Salviagnol, I.; Blancher, A.; Calvas, P.; Clayton, J.; Socha, W.W.; Colln, Y.; Ruffe,
Biochem. Genet. 32, 201-221, 1994
A:Title: Molecular genetics of chimpanzee Rh-related genes: their relationship with t
A:Reference number: I37003; MUID:95085395
A:Accession: I37004
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-417 <RES>
A:Cross-references: GB:I37049; NID:g606989; PIDN:AAA65623.1; PID:g606990

Query Match 86.6%; Score 1840; DB 2; Length 417;
Best Local Similarity 88.7%; Pred. No. 1.8e-141;
Matches 37/0; Conservative 11; Mismatches 36; Indels 0; Gaps 0;

```

1 MSCKYPRSVGRCLPLCALTEALILLFYFTHYDASLEDDKGLVASYOVGODLTVAAI 60
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1 MSCKYPRSVGRCLPLCALTEALILLFYFTHYDASLEDDKGLVASYOVGODLTVAAI 60
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 61 GLGFLTSSFRHSSWSVAFLFMLALGVOMAILDGLFLOSPPSGKVITTFISILATMSA 120
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 61 GFGFLTSSFRHSSWSVAFLFMLALGVOMAILDGLFLOSPPSGKVITTFISILATMSA 120
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 121 LSVLISDAVLGKYNLAQLVVMVLEVTDLGNLRMVISNIFNTDYHNNMMHIYFAAYFG 180
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 121 LSVLISDAVLGKYNLAQLVVMVLEVTDLGNLRMVISNIFNTDYHNNMMHIYFAAYFG 180
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 181 LTVAMCLPKPLPECTEDNDQATIPSLSAMLGALFLMFPSPVSALLRSPIERKNVFN 240
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 181 LTVAMCLPKPLPECTEDNDQATIPSLSAMLGALFLMFPSPVSALLRSPIERKNVFN 240
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 241 TYAVAVSVVTAISGSSSLAHPOGKISKTYGHSVAVLPEGVAVDTSCHLIPSPWLPVLGLV 300
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 241 TYAVAVSVVTAISGSSSLAHPOGKISKTYGHSVAVLPEGVAVDTSCHLIPSPWLPVLGLV 300
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 301 AGLISVYGAKYLPCCNRRVLGIPHSSTMGYNFSLGLLEIITYVLVLDTVGANGMIG 360
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 301 AGLISVYGAKYLPCCNRRVLGIPHSSTMGYNFSLGLLEIITYVLVLDTVGANGMIG 360
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 361 FOVLSTIGELSLAIVLTGSLTGLTALLNKIRKAPHEAKYFDQVFWKPEHLAVGF 417
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 361 FOVLSTIGELSLAIVLTGSLTGLTALLNKIRKAPHEAKYFDQVFWKPEHLAVGF 417
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

RESULT 6

A30405
erythrocyte membrane protein Rh (unknown specificity) - human
N:Alternate names: blood group Rh antigen
C:Species: Homo sapiens (man)
C>Date: 05-Oct-1990 #sequence_revision 05-Oct-1990 #text_change 21-Jul-2000
C/Accession: A30405; S13060; I34273; S40515; S40516; S40517; S02085; S02087; A32509;
R:Cherif-Zahar, B.; Bloy, C.; Le Van Kim, C.; Blanchard, D.; Bailly, P.; Hermand, P.;
Proc. Natl. Acad. Sci. U.S.A. 87, 6243-6247, 1990


```

QY 241 TYVAAVSVTAISGSSLAHPGOKISKTYGHSVAVLPEGVADTSCHLIPSPWLDIVGLV 300
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 241 TYVALAVSVTAISGSSLAHPGOKISMTYVHSAVLAGVANGTSCHLIPSPWLAIVGLV 300

QY 301 AGLISVCGAKYLPCCCKNRVIGIPHSSTMGVNFSLGLLEIITYVLVDIVGANGMIG 360
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 301 AGLISGACAGCCCKNRVIGIPDSVHYNFSLGLLEIITYVLVLRHTVANGMIG 360

QY 361 FOVLISGELSLAIVALTSGTLTALLNLRKAPHAKEFDDQVFWKPFHLAVGF 417
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 361 FOVLISGELSLAIVALTSGTLTALLNLRKAPHAKEFDDQVFWKPFHLAVGF 417

```

RESULT 8

```

137075
Rhesus-like protein - gorilla
C:Species: Gorilla gorilla (gorilla)
C>Date: 04-Oct-1996 #sequence_revision 04-Oct-1996 #text_change 21-Jul-2000
C:Accession: I37075
R:Salvignol, I.; Blancher, A.; Calvas, P.; Clayton, J.; Socha, W.W.; Collin, Y.; Ruffie,
Biochem. Genet. 32, 201-221, 1994
A:Title: Molecular genetics of chimpanzee Rh-related genes: their relationship with the
A:Reference number: I37003; MUID:95085595
A:Accession: I37075
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-417 <RES>
A:Cross-references: GB:I37052; NID:g607009; PIDN:AAA65626.1; PID:g607010

```

```

Query Match      84.8%; Score 1803; DB 2; Length 417;
Best Local Similarity 86.6%; Pred. No. 1.7e-138;
Matches 361; Conservative 15; Mismatches 41; Indels 0; Gaps 0;

```

```

QY 1 MSCRYPSRVGRCPLCALITLLEAAILLFFFTYHDASLEDDKGLVASYOGODLTVMMAI 60
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1 MSSKYPSPVRCCLPLCALITLLEAAILLFFFTYHDASLEDDKGLVASYOGODLTVMMAI 60

QY 61 GLGFLTSPFRHSSVSAFNFMLALGVQMAILLDFISQPSKVVITLFSIALATMSA 120
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 61 GFGFLTSPFRHSSVSAFNFMLALGVQMAILLDFISQPSKVVITLFSIALATMSA 120

QY 121 LSVLISVDVAVLGNVLAQLVVMVLEVTDLGNLRMVISNIFNTDYHMMMHIVYFAAYFG 180
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 121 LSVLISGAVLGVNIQVLMVLEVTALGTMMVSNIFNTDYHMMMHIFVYFAAYFG 180

QY 181 LTVAMCLPKPLPEGTEDNDQATIPSLSAMUGALFLWMEFVSNSALIRSPIERKNAYFN 240
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 181 VTVAMCLPKPLPEGTEDNDQATIPSLSAMUGALFLWMEFVSNSALIRSPIERKNAYFN 240

QY 241 TYVAAVSVTAISGSSLAHPGOKISKTYGHSVAVLPEGVADTSCHLIPSPWLDIVGLV 300
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 241 TYVALAVSVTAISGSSLAHPGOKISMTYVHSAVLAGVANGTSCHLIPSPWLAIVGLV 300

QY 301 AGLISVCGAKYLPCCCKNRVIGIPHSSTMGVNFSLGLLEIITYVLVDIVGANGMIG 360
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 301 AGLISGACAGCCCKNRVIGIPDSVHYNFSLGLLEIITYVLVLRHTVANGMIG 360

QY 361 FOVLISGELSLAIVALTSGTLTALLNLRKAPHAKEFDDQVFWKPFHLAVGF 417
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 361 FOVLISGELSLAIVALTSGTLTALLNLRKAPHAKEFDDQVFWKPFHLAVGF 417

```

RESULT 9

```

184434
Rhesus-like protein - crab-eating macaque
C:Species: Macaca fascicularis (crab-eating macaque)
C>Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 21-Jul-2000
C:Accession: I84434
R:Salvignol, I.; Blancher, A.; Calvas, P.; Clayton, J.; Socha, W.W.; Collin, Y.; Ruffie,
Biochem. Genet. 32, 201-221, 1994
A:Title: Molecular genetics of chimpanzee Rh-related genes: their relationship with the
A:Reference number: I37003; MUID:95085595

```

```

A:Accession: I84434
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-417 <RES>
A:Cross-references: GB:I37054; NID:g609515; PIDN:AAA65626.1; PID:g609516

```

```

Query Match      71.0%; Score 1509; DB 2; Length 417;
Best Local Similarity 74.3%; Pred. No. 1.1e-114;
Matches 310; Conservative 31; Mismatches 76; Indels 0; Gaps 0;

```

```

QY 1 MSCRYPSRVGRCPLCALITLLEAAILLFFFTYHDASLEDDKGLVASYOGODLTVMMAI 60
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1 MSSKYPSPVRCCLPLMALITLLEAAILLFFFTYHDASLEDDKGLVASYOGODLTVMMAI 60

QY 61 GLGFLTSPFRHSSVSAFNFMLALGVQMAILLDFISQPSKVVITLFSIALATMSA 120
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 61 GLGFLTSPFRHSSVSAFNFMLALGVQMAILLDFISQPSKVVITLFSIALATMSA 120

QY 121 LSVLISVDVAVLGNVLAQLVVMVLEVTDLGNLRMVISNIFNTDYHMMMHIVYFAAYFG 180
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 121 TSMLSMNAVLAGKNVLAQLVVMVLEVTFTGTMKRVITNFKIDYGNMMHIVFAAYFG 180

QY 181 LTVAMCLPKPLPEGTEDNDQATIPSLSAMUGALFLWMEFVSNSALIRSPIERKNAYFN 240
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 181 LTVAMCLPKPLPKGTEDKDYQTTSPLFAMIGTLFLWMEFVSNSALIRSPIERKNAYFN 240

QY 241 TYVAAVSVTAISGSSLAHPGOKISKTYGHSVAVLPEGVADTSCHLIPSPWLDIVGLV 300
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 241 TYVALAVSVTAISGSSLAHPGOKISMTYVHSAVLAGVANGTSCHLIPSPWLAIVGLV 300

QY 301 AGLISVCGAKYLPCCCKNRVIGIPHSSTMGVNFSLGLLEIITYVLVDIVGANGMIG 360
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 301 AGLISFGACAGCCCKNRVIGIPDSVHYNFSLGLLEIITYVLVLRHTVANGMIG 360

QY 361 FOVLISGELSLAIVALTSGTLTALLNLRKAPHAKEFDDQVFWKPFHLAVGF 417
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 361 FOVLISGELSLAIVALTSGTLTALLNLRKAPHAKEFDDQVFWKPFHLAVGF 417

```

RESULT 10

```

578480
Rhesus blood group antigen-like protein isoform 4 - human
C:Species: Homo sapiens (man)
C>Date: 11-Mar-1998 #sequence_revision 17-Apr-1998 #text_change 05-Nov-1999
C:Accession: S78480; S26560
R:Collin, Y.
submitted to the EMBL Data Library, November 1991
A:Reference number: S78478
A:Accession: S78480
A:Molecule type: mRNA
A:Residues: 1-354 <COL>
R:Le Van Kim, C.; Cherif-Zahar, B.; Reynal, V.; Mourou, I.; Lopez, M.; Cartron, J.P.;
Blood 80, 1074-1078, 1992
A:Title: Multiple Rh messenger RNA isoforms are produced by alternative splicing.
A:Reference number: S26560; MUID:92360855
A:Accession: S26560
A:Molecule type: mRNA
A:Residues: 160-354 <LEV>
A:Cross-references: EMBL:X63098
A:Experimental source: tissue-type bone marrow
C:Keywords: alternative splicing; erythrocyte; surface antigen; transmembrane protein

```

```

Query Match      66.4%; Score 1411; DB 2; Length 354;
Best Local Similarity 90.1%; Pred. No. 8e-107;
Matches 282; Conservative 10; Mismatches 21; Indels 0; Gaps 0;

```

```

QY 1 MSCRYPSRVGRCPLCALITLLEAAILLFFFTYHDASLEDDKGLVASYOGODLTVMMAI 60
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1 MSSKYPSPVRCCLPLCALITLLEAAILLFFFTYHDASLEDDKGLVASYOGODLTVMMAI 60

```

Qy	61	GLGFIETSEFRHSHMSVAFNLFMLALGVOMAILLDGFLSOPSGKVVITLFSLIWMATSA	120
Db	61	CLGFIETSEFRHSHMSVAFNLFMLALGVOMAILLDGFLSOPSGKVVITLFSLIWMATSA	120
Qy	121	LSVLISVDALGVKYNLAOLVYVVLVEVYTDGNIIRMVISNIFNTDQYHNMKHIYVFAAFG	180
Db	121	MSVLISACAVAGIKYNLAOLVYVVLVEVYTAGTLIRMTVISNIFNTDQYHNMKHIYVFAAFG	180
Qy	181	LTVMACLPKPLPEGTEDDDQATIPISISAMIGALFIMFPPSVNSALLBSPIERKNAVEN	240
Db	181	LTVMACLPKPLPEGTEDDDQATIPISISAMIGALFLMFWPSSVNSALLRSPIDRKNAFEN	240
Qy	241	TYVAANAVSVTAISGSSLIAPHQGIKSTYGHSAVLPEGVAVDTSCHLIPSPMDIPILVGLV	300
Db	241	TYVAANAVSVTAISGSSLIAPHQGIKSTYGHSAVLPEGVAVDTSCHLIPSPMDIPILVGLV	300
Qy	301	AGLISVYGAKLP 313	
Db	301	AGLISVYGAKLP 313	

RESULT 11
I37053
Rhesus-like protein - pileated gibbon
C:Species: Hylobates pileatus (Pileated gibbon)
C:Date: 04-Sep-1997 #sequence_revision 04-Sep-1997 #text_revision 05-Nov-1999
C:Accession: I37053
R:Salviagnol, I.; Blanchet, A.; Calvas, P.; Clayton, J.; Socha, W.W.; Colin, Y.; Ruffie,
Biochem. Genet. 32, 201-221, 1994
A:Title: Molecular genetics of chimpanzee Rh-related genes: their relationship with the
A:Reference number: I37003; MUID:95085595
A:Accession: I37053
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-354 <RES>
A:Cross-references: GB:I37051; NID:.9607007; PIDN:AAA5625.1; PID:9607008

	Query Match	Similarity	60.9%	Score 1294	DB 2	Length 354	
	Best Local	Similarity	82.7%	Pred. No. 2.4e-97			
	Matches	259	Conservative	20	Mismatches	34	Indels
							Gaps
QY	1	MSCKYRSVAGRCPLCATLLEALAILLEFFPHYASLSDQGLVASYGVODLTVMMAI	60				
Db	1	MSSKYRSVAGRCPLMALTLLEAILLLEFFPHYDASLSDQGLVATVGVODLTVMMAI	60				
QY	61	GIGFTTSFRRHSWSSVAFNLFMLALGYOMAILDGFISQPPSGKVYITLSITLATMSA	120				
Db	61	GIGFTLSNLRHSWSSVAFNLFMLALGYOMAILDGFISQPPSGKVYITLSIRLATMSA	120				
QY	121	LSVLISVAVAGKVAADLVVAVLVLEVVDLGLNLRVVISIFMTDVMNMNHLYVFAAYSG	180				
Db	121	MSSLISVGVAGKVAADLVVAVLVLELTPFGTRRVINNIFFTDVHVMNHVFPAAVFG	180				
QY	181	LTVAACLPKPLPECTEDNDQFATISLSAMLCALFLMMPSPVNSALRSPIERKNAVEN	240				
Db	181	LTVAACLPKPRPEGIEDDEQGLANSLSAMLCALFLMIFWPFNSALLTNPILERKNAVEN	240				
QY	241	TYVAVAVSVTAIAGSSSLAHPOGKISKITYGHSAVLDPGVAVDTSCHLIPSPWLPIVLGLV	300				
Db	241	TYVAVAVSVTAIAIVSSSLAHPOGKINMTYMHNAVLAVGVAAGVTSCHLISSPWLAVLGLV	300				
QY	301	AGLISVTCAGKILP 313					
Db	301	AGLISIGGAKCLP 313					

RESULT 12
S78479 Rhesus blood group antigen-like protein isoform VIII - human
C:Species: Homo sapiens (man)
C:Date: 11-Mar-1998 #sequence_revision 17-Apr-1998 #text_change 05-Nov-1998
#Accession: S78479; S28563

R:Collin, Y.
submitted to the EMBL Data Library, November 1991
A:Reference number: S78478
A:Accession: S78479
A:Molecule type: mRNA
A:Residues: 1-266 <COL>
A:Cross-references: NID:X63096; NID:g36044; PIDN:CA44810.1; PID:g36045
R:Le Van Kim, C.; Cheifetz-Zahar, B.; Reynal, V.; Mourou, I.; Lopez, M.; Carlton, J.P.;
Biolld 80, 1074-1078, 1992
A:Title: Multiple Rh messenger RNA isoforms are produced by alternative splicing.
A:Reference number: S26560; MUID:92360855
A:Accession: S26563
A:Molecule type: mRNA
A:Residues: 160-266 <LEV>
A:Cross-references: EMBL:X63096
A:Experimental source: tissue-type bone marrow
C:Keywords: alternative splicing; erythrocyte; surface antigen; transmembrane protein

Query Match	51.3%	Score 1089.5	DB 2	Length 266
Best Local Similarity	58.0%	Pred. No. 6,2e-81		
Matches 242	Conservative 3	Mismatches 21	Indels 151	Gaps
QY	1	MSCKYPSRVGRCPLCATLTTEAALILLFYFFTHYDASIEDQKGLVASYQVGDLTGMAAI	60	
		: : : : : : : : :		
DB	1	MSKKPRSVRRCRLPCATLTTEAALILLFYFFTHYDASIEDQKGLVASYQVGDLTGMAAI	60	
QY	61	GLGLTTSFRHRHSSVAFNLFMLALGVQAAIILDDGLSQPSQSKVYITLFTSLATMSA	120	
		: : : : : : : : :		
DB	61	GLGLTTSFRHRHSSVAFNLFMLALGVQAAIILDDGLSQPSQSKVYITLFTSLATMSA	120	
QY	121	LSVLISVAVLVGKNIQLQVLVYVLEVTDLGNLRMVFSNIFNFDYHNMHMIYVFAAFG	180	
		: : : : : : : : :		
DB	121	MSVLISAAVLGKNIQLQVLVYVLEVTALGTLEKMTVSNIFNV-----	163	
QY	181	LTVAWCLPKPLPEGTEDNDQATIPSLSMALGALFLMFT-PSVNSALLRSPIERKKAIVEN	240	
		: : : : : : : : :		
DB	164	-----	163	
QY	241	TYANAASVVTALISGSSLAHPOGKISKTYGHSAAVLPGGVAVDTSCHLIPSPMLPIVLGLV	300	
		: : : : : : : : :		
DB	164	-----	163	
QY	301	AGLISVIGAKYLLPGCCNKNRVLGIPHSIMGVNFSLGLLEETIYVLLVDITVGAGNGMIG	360	
		: : : : : : : : :		
DB	164	-----CNRVLGIHHISVMHSIFSLGLLEETIYVLLVDITVWNGNGMIG	209	
QY	361	FOYLLISIGELSIAIVIALTSSLLTALLNLKLTAAHREAKYFPDDQVYWKRPRLHAYG	417	
		: : : : : : : : :		
DB	210	FOYLLISIGELSIAIVIALTSSLLGGLLNLIKIMKAPAAVAYFPDDQVYWKRPRLHAYG	266	

RESULT 13
S78478
Rhesus blood group antigen-like protein isoform VI - human
C/Species: Homo sapiens (man)
C/Date: 11-Mar-1998 #sequence_revision 17-Apr-1998 #text_change 05-Nov-1999
C/Accession: S78478; S26562
R/Colin, Y.
submitted to the EMBL Data Library, November 1991
A/Reference number: S78478
A/Accession: S78478
A/Molecule type: mRNA
A/Residues: 1-267 <COL>
A/Cross-references: EMBL:X63095; NID:g36042; PIDN:CAA44809.1; PID:g36043
R/Le Van Kim, C.; Cherif-Tahar, B.; Raynal, V.; Mouru, I.; Lopez, M.; Cartron, J.P...
Blood 80, 1074-1078, 1992
A/Title: Multiple Rh messenger RNA isoforms are produced by alternative splicing.
A/Reference number: S26560; MUID:92360855
A/Accession: S26562
A/Molecule type: mRNA
A/Residues: 160-267 <LEV>
A/Cross-references: EMBL:X63095

A: Experimental source: tissue-type bone marrow
C: Keywords: alternative splicing; erythrocyte; surface antigen; transmembrane protein

Query Match	47.1%;	Score 1000.5;	DB 2;	Length 267;
Best Local Similarity	61.5%;	Pred. No. 1e-73;		
Matches 220;	Conservative 7;	Mismatches 26;	Indels 105;	Gaps 1;

QY	1	MSCKPPSVSAGRCPLCALTLTEAAILLFYFFFTHYDASLEQKVLASVYQGDGLTYMAAI	60
		: : : : : : : : : : : :	
Db	1	MSSTPPSVSRRCPLMALTEEAAILLFYFFFTHYDASLEQKVLASVYQGDGLTYMAAL	60
QY	61	GLGFLTSSFRHSSMSVAFNLFMLALGVOMAILLDGLSQPSGKVITLFSLTATMSA	120
		: : : : : : : : : : :	
Db	61	GLGFLTISNFRHSSMSVAFNLFMLALGVOMAILLDGLSQPPGKVITLFSIRLTMSA	120
QY	121	LSVTLSDVAVLGKKNLALQVLVYMLVEVETDGLGNLRMYSNIFENFDYHMMNHIVFAAYG	180
		: : : : : : : : : : :	
Db	121	MSVTLISAGVVLGKKNLALQVLVYMLVEVETALGTLRMVINSIFN-----	162
QY	181	LTVAACLPRKPLPEGTEDNDQRTATPLSLAMLAGLFIMGF-PSVNSALILRSPIERKNAVEN	240
Db	163	-----	162
QY	241	TYTVAVNSVYVTAISGSSLAPDCKISTFYHSVAVLPRGVAVDTSCHLIPSPWLPJVLGLV	300
		: : : : : : : : : : :	
Db	163	-----TYHSVAVLGAVAVGTSCHLIPSPWLAIVGLV	195
QY	301	AGLISY-GAKYLPGCCNRRVLGIPHSITMGVNFSLILGLLEIITVILVLDYVAGAGGM	358
		: : : : : : : : : : :	
Db	196	AGLISYGAKCLPVCCKNRRVLGIIHSIVMSHIFSLILGLLEIITVILVLDYVAVNGNGM	253

RESULT 14

membrane glycoprotein, 50K, erythrocyte - human
C:Species: Homo sapiens (man)
C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 01-Dec-2000
C:Accession: S29124; JEO168
R:Ridgwell, K.; Spurr, N.K.; Laguda, B.; Macgeoch, C.; Avent, N.D.; Tanner, M.J.A.
Biochem. J. 287, 223-228, 1992
A:Title: Isolation of cDNA clones for a 50 kDa glycoprotein of the human erythrocyte mem
A:Reference number: S29124; MUID:93038538
A:Accession: S29124
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-409 <RID>
A:Cross-references: EMBL:X64594; NID:g31194; PIDN:CAA45883.1; PID:g31195
R:Iwamoto, S.; Omi, T.; Yamasaki, M.; Okuda, H.; Kawano, M.; Kajii, E.
Biochem. Biophys. Res. Commun. 243, 233-240, 1998
A:Title: Identification of 5' flanking sequence of RH50 gene and the core region for eryth
A:Reference number: JEO168; MUID:98139897
A:Accession: JEO168
A:Molecule type: mRNA
A:Residues: 1-52 <IWA>
C:Comment: This protein is membrane protein tightly associated with Rh polypeptides and
C:Keywords: glycoprotein; transmembrane protein
A:Gene: RH50
C:Genetics:

Query Match 25.8%; Score 547.5; DB 2; Length 409;
Best Local Similarity 33.9%; Pred. No. 8e-37;
Matches 139; Conservative 79; Mismatches 171; Indels 21; Gaps 77

QY 14 PLCAITAEALILFEFFETH--PASHIEDOKGLVASQV-----ODLTVMAAIGL 62
| | | | | | | | | | | | | | | | | | | | | | | | | |
Db 6 PLMAIVEIAIVTIFGFEVEETDQTVEIQNTIKPPDMGTFEELIYPEDDVHAIIVEGF 65
| | | | | | | | | | | | | | | | | | | | | | | | | |
QY 63 GLTISSFRHSWSVAENLEMALGVQWAILLDGFLSQDPBGKVVITFSLWATMSALS 122
| | | | | | | | | | | | | | | | | | | | | | | | | |
Db 66 GLMPLFKKKYCFSSGINLLVALAGLTGMGTVOGIT--OSOGKRNIETIKNNINDFSNAAT 124
| | | | | | | | | | | | | | | | | | | | | | | | | |

[illegible]

RESULT 15

Hypothetical protein F08F3.3 - *Caenorhabditis elegans*
 C.Species: *Caenorhabditis elegans*
 C.Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
 C.Accession: J29442
 R.Blanchard, M.; Bradshaw, H.
 submitted to the EMBL Data Library, July 1996
 A.Description: The sequence of C. elegans cosmid F08F3.
 A.Reference number: Z20620
 A.Accession: J29442
 A.Status: preliminary; translated from GB/EMBL/DDbJ
 A.Molecule type: DNA
 A.Residues: 1-463 <BLA>
 A.Cross-references: EMBL:U64847; PIDD:AA04874.1; GSPDB:GN00023; CESP:F08F3.3
 A.Experimental source: strain Bristol N2; clone F08F3
 C.Genetics:
 A.Gene: CESP:F08F3.3
 A.Map position: 5
 A.Introns: 18/3; 167/3; 266/3; 362/2; 442/3

Query Match 19.4%; Score 412; DB 2; Length 463;
Best Local Similarity 26.6%; Pred. No. 8.7e-26;
Matches 116; Conservative 88; Mismatches 190; Indels 42; Gaps 8,

```

0Y 10 qRcPLcALTLtEALlLLfYFFtHNDs-----LEdQGLVAsYQVGDdLTVMAIG 61
   | | | | | : : : | : | | | |
Db 7 QNOLTLGLfQVfVfALfYALGSYdASALPSEtKNKEEAARMLNLPfLQDTHVMfIG 66
   | | | | | : : : | : | | | |
0Y 62 LGfLTSFRfRHSMSVAVNLfMLALGYOMALLDGLfLSQfRPSGKVTLfTSfVLTATMSAL 121
   | | | : : : | : | : : : | : | : : : | : | : : : | : | : : : |
Db 67 fEFLTLfLKRfGSfASVSLNMLLAfFTLQWGLTVfRGMAShHGEKfFTLSLEOLLTADfAA 128
   | | | : : : | : | : : : | : | : : : | : | : : : | : | : : : |
0Y 122 SVLISVDVdLVKVNLAQVVMVLEVTLdGLMfRWYSLNfFTDfYHNM-----MMHfYFAA 177
   | | | : : | : | : | : | : : : | : : : : : | : : : : : |
Db 127 VILISMGMfLKLfSfQVYfNAfFTP-----VALfVEHfTCYHNfQINDVGSSTfVAFGA 182
   | | | : : | : | : | : | : : : | : : : : : | : : : : : |
0Y 178 YfGLfVAMCLPKPLfRfEGTfEDDQfATfPISLSAMfLGAfLTMfRfPSVNSALfRSPfIERKNA 237
   | | | | | : | : | : | : | : | : | : | : | : | : | : | : |
Db 183 YfGLfCANGfGfKKfERG--HTfMGfSTfYfHfDfIffAMfGALfPMTfYfMfSPfAAVAATDfAQR 244
   | | | | | : | : | : | : | : | : | : | : | : | : | : | : |
0Y 238 VENTfYAAVAVSVTAfSSSLAHfRQGLSKfTYfGHSAVLfPecAVAdTSCHfLIPfMPLfVL 297
   | | | : : | : | : | : | : : | : | : | : | : | : | : | : |
Db 242 VANfFLSLfCfACfMTfELVfSQAVdKHfKRfRfDfWHfANfSTLAGVAlGTfANYVLEBYHAMfII 301
   | | | : : | : | : | : | : : | : | : | : | : | : | : | : |
0Y 298 GLVAGLISVfGAKYVLPGCCfNRVLGfPHSSfLNGfVfNSfLGLdELfIYVLVL-----3q9
   | | | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 302 GVfAGAVSVfVfGKYfTfPLfSKLdHfDfCGVNNLHGMPGLfAGfPASfIAfLfYfDEfTRPA 361
   | | | : | : | : | : | : | : | : | : | : | : | : | : | : |
0Y 350 --dYV--GAGNG-----MGfOVfLLSfIGELSLAfYALfTSGfLLTfALLfNLKfITKA 395
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : |

```

Db 362 QYDKIYPGMARGEDRTRMFDEKTOALNQLMAIGLVFLASTV---SGYLTGILLKLIKIMDQ 418
QY 396 PHEAKYFDDQYFWKFP 411
 : :1: | : : : |
Db 419 VRDDEYYADGSDYFETP 434

Search completed: September 12, 2001, 07:39:35
Job time: 159 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 12, 2001, 07:39:11 ; Search time 9.37 Seconds

(without alignments)
1528.153 Million cell updates/sec

Title: 09-600714-2

Perfect score: 2125

Sequence: 1 MSCKYPSVGRCLPLCALTL.....AKYDDQVFWKPEHLAVGEF 418

Scoring table: BIOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 93435 seqs, 34255486 residues

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : SwissProt_39.*

Listing first 45 summaries

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1975	92.9	416	1 RHD_HUMAN	002161 homo sapien
2	1858	87.4	416	1 RHLP_PANTR	028814 pan troglod
3	1837	86.4	416	1 RHLD_GORGO	028427 gorilla gor
4	1835	86.4	416	1 RHLP_PANTR	028813 pan troglod
5	1828	86.0	416	1 RHCE_HUMAN	028817 homo sapien
6	1806	85.0	416	1 RHLP_PANTR	028812 pan troglod
7	1798	84.6	416	1 RHLC_GORGO	028426 gorilla gor
8	1519	71.5	416	1 RHLP_PANTR	028814 pan troglod
9	1504	70.8	416	1 RHLP_PANTR	028814 pan troglod
10	1289	60.7	416	1 RHLP_PANTR	028814 pan troglod
11	547.5	25.8	409	1 RHAG_HUMAN	002094 homo sapien
12	119.5	5.6	388	1 RHLP_PANTR	028814 pan troglod
13	112.5	5.3	416	1 RHLP_PANTR	028814 pan troglod
14	111.5	5.2	416	1 RHLP_PANTR	028814 pan troglod
15	111.5	5.2	416	1 RHLP_PANTR	028814 pan troglod
16	111.5	5.2	416	1 RHLP_PANTR	028814 pan troglod
17	109.5	5.2	416	1 RHLP_PANTR	028814 pan troglod
18	109.5	5.2	416	1 RHLP_PANTR	028814 pan troglod
19	109.5	5.2	416	1 RHLP_PANTR	028814 pan troglod
20	107.5	5.1	416	1 RHLP_PANTR	028814 pan troglod
21	107.5	5.1	416	1 RHLP_PANTR	028814 pan troglod
22	107.5	5.1	416	1 RHLP_PANTR	028814 pan troglod
23	106.5	5.0	416	1 RHLP_PANTR	028814 pan troglod
24	106.5	5.0	416	1 RHLP_PANTR	028814 pan troglod
25	106.5	5.0	416	1 RHLP_PANTR	028814 pan troglod
26	106.5	5.0	416	1 RHLP_PANTR	028814 pan troglod
27	103.5	4.9	416	1 RHLP_PANTR	028814 pan troglod
28	103.5	4.9	416	1 RHLP_PANTR	028814 pan troglod
29	103.5	4.9	416	1 RHLP_PANTR	028814 pan troglod
30	102.5	4.8	416	1 RHLP_PANTR	028814 pan troglod
31	102.5	4.8	416	1 RHLP_PANTR	028814 pan troglod
32	100.5	4.7	416	1 RHLP_PANTR	028814 pan troglod
33	100.5	4.7	416	1 RHLP_PANTR	028814 pan troglod

34	100	4.7	416	1 YBDA_ECOLI	P24077 escherichia
35	100	4.7	452	1 NU4M_BRALA	074423 branchiosto
36	100	4.7	452	1 NU4M_BRALA	079421 branchiosto
37	100	4.7	887	1 Y277_MYCPN	P75387 mycoplasma
38	100	4.7	982	1 YS96_CABEL	009965 caenorhabd
39	99.5	4.7	594	1 COX1_PHYPO	007434 physarum po
40	99	4.7	603	1 NU5M_PANPA	P03916 pan paniscu
41	98.5	4.6	448	1 GNTF_BACSU	P12012 bacillus su
42	98.5	4.6	509	1 NUOM_ECOLI	P31978 escherichia
43	98	4.6	338	1 YFJY_ECOLI	P52137 escherichia
44	98	4.6	347	1 NU2M_DIDMA	P41305 didelphis m
45	98	4.6	442	1 YCDE_ECOLI	P75892 escherichia

ALIGNMENTS

RESULT	ID	STANDARD	PRT	AA
1	RHD_HUMAN	002161; 007618; 016355; 016147; 016235;		
AC	002161; 007618; 016355; 016147; 016235;			
DT	01-NOV-1997 (Rel. 35, Created)			
DT	01-NOV-1997 (Rel. 35, Last sequence update)			
DT	01-OCT-2000 (Rel. 40, Last annotation update)			
DE	BLOOD GROUP RH(D) POLYPEPTIDE (RHESUS D ANTIGEN) (RHXXIII) (RH			
DE	POLYPEPTIDE 2) (RHPHII).			
GN	RHD.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Bone marrow;			
RX	MEDLINE=9306356; PubMed=1438298;			
RA	le van Kim C., Mouro I., Cherrier C.,			
RA	Carton J.-P., Collin Y.,			
RT	"Molecular cloning and primary structure of the human blood group Rhd			
RT	polypeptide.";			
RL	Proc. Natl. Acad. Sci. U.S.A. 89:10925-10929(1992).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Bone marrow;			
RX	MEDLINE=9320449; PubMed=8329718;			
RA	Arce M.A., Thompson E.S., Wagner S., Coyne K.E., Ferdman B.A.,			
RA	Lublin D.M.,			
RT	"Molecular cloning of Rhd cDNA derived from a gene present in Rhd-			
RT	positive, but not Rhd-negative individuals.";			
RL	Blood 82:651-655(1993).			
RN	[4]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=93216282; PubMed=7916743;			
RA	Kajil E., Umenishi F., Iwamoto S., Ikemoto S.,			
RT	"Isolation of a new cDNA clone encoding an Rh polypeptide associated			
RT	with the Rh blood group system.";			
RL	Hum. Genet. 91:157-162(1993).			
RN	[5]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=95329738; PubMed=7606008;			
RA	Huang C.H., Reid M.E., Chen Y.,			
RT	"Identification of a partial internal deletion in the RH locus			
RT	causing the human erythrocyte D-phenotype.";			
RL	Blood 86:784-790(1995).			
RN	[6]			
RP	SEQUENCE FROM N.A. (SHORT FORM 1).			

RX MEDLINE=94235883; PubMed=8180407;
RA Westhoff C.M., Wylie D.E.;
RT "Identification of a new RhD-specific mRNA from K562 cells.";
RL Blood 83:3098-3100(1994).
RP [7]
RX SEQUENCE FROM N.A. (SHORT FORM 2).
RX MEDLINE=94362249; PubMed=8080999;
RA Suyama K., Lunn R., Haller S., Goldstein J.;
RT "Rh(D) antigen expression and isolation of a new Rh(D) cDNA isoform
in human erythroleukemic K562 cells.";
RL Blood 84:1975-1981(1994).
RN [8]
RP VARIANT BLOOD GROUP TAR.
RX MEDLINE=95259709; PubMed=7741145;
RA Rouillac C., le van Kim C., Beolet M., Cartion J.-P., Colin Y.;
RT "Leulopro substitution in the RhD polypeptide is responsible for the
DVI1 category blood group phenotype.";
RL Am. J. Hematol. 49:87-88(1995).
CC -1- FUNCTION: MAY BE PART OF AN OLIGOMERIC COMPLEX WHICH IS LIKELY TO
HAVE A TRANSPORT OR CHANNEL FUNCTION IN THE ERYTHROCYTE MEMBRANE.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -1- ALTERNATIVE PRODUCTS: VARIOUS FORMS ARE PRODUCED BY ALTERNATIVE
SPlicing.
CC -1- TISSUE SPECIFICITY: RESTRICTED TO TISSUES OR CELL LINES EXPRESSING
ERYTHROID CHARACTERS.
CC -1- POLYMORPHISM: RHD AND RHCE ARE RESPONSIBLE FOR THE RH BLOOD GROUP
SYSTEM. THE MOLECULAR BASIS OF THE TAR-RH40 BLOOD GROUP ANTIGEN
IS A POLYMORPHISM IN POSITION 109.
CC -1- SIMILARITY: BELONGS TO THE RH FAMILY. STRONG, TO RHCE.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X63097; CAA44811.1; -
DR EMBL: X63094; CAA44808.1; -
DR EMBL: L08429; AAA02679.1; -
DR EMBL: S57971; AAB36081.1; -
DR EMBL: S78509; AAB34852.1; -
DR EMBL: S70174; AAB30756.1; -
DR EMBL: S73913; AAB31911.1; -
DR PIR: S26564; S26564.
DR MIM: 111680; -
DR InterPro: IPR001905; -
DR InterPro: IPR002229; -
DR Pfam: PF00909; Ammonium_transp; 1.
DR PRINTS: PR00342; RHESUSRHD.
KM Erythrocyte; Transmembrane; Blood group antigen; Alternative splicing;
KW Polymorphism.
FT INIT_MET 0
FT TRANSMEM 11 31 POTENTIAL.
FT TRANSMEM 43 63 POTENTIAL.
FT TRANSMEM 76 96 POTENTIAL.
FT TRANSMEM 106 126 POTENTIAL.
FT TRANSMEM 129 149 POTENTIAL.
FT TRANSMEM 166 186 POTENTIAL.
FT TRANSMEM 202 222 POTENTIAL.
FT TRANSMEM 237 257 POTENTIAL.
FT TRANSMEM 286 306 POTENTIAL.
FT TRANSMEM 333 353 POTENTIAL.
FT TRANSMEM 357 377 POTENTIAL.
FT VARSPPLIC 313 408 MISSING (IN SHORT ISOFORM 1).
FT VARSPPLIC 315 415 C -> S (IN SHORT ISOFORM 2).
FT VARSPPLIC 316 416 MISSING (IN SHORT ISOFORM 2).
FT VARIANT 109 109 L -> P (IN TAR ANTIGEN).
FT VARIANT 109 109 /FTID-VAR_006919.
FT VARIANT 217 217 M -> I.
FT VARIANT 217 217 /FTID-VAR_006920.
FT CONFLICT 15 15 W -> C (IN REF. 5).
CC CONFLICT

FT CONFLICT 38 38 E -> G (IN REF. 4).
FT CONFLICT 102 102 S -> P (IN REF. 4).
FT CONFLICT 126 126 V -> A (IN REF. 4).
FT CONFLICT 173 173 V -> M (IN REF. 5).
FT CONFLICT 173 173 S -> T (IN REF. 4).
FT CONFLICT 181 181 S -> V (IN REF. 4 AND 7).
FT CONFLICT 313 313 G -> T (IN REF. 4).
FT CONFLICT 322 322 P -> H (IN REF. 4).
FT CONFLICT 397 397 E -> V (IN REF. 5).
SQ SEQUENCE 416 AA; 45049 MM; 9136DPLA37D7681B CRC64;
Query Match 92.9%; Score 1975; DB 1; Length 416;
Best Local Similarity 95.2%; Pred. No. 1.2e-136;
Matches 396; Conservative 3; Mismatches 17; Indels 0; Gaps 0;
QY 2 SCKYPRVQGLPLCATLLEAALLLFYFTTHDASLEDDKGLVASYQVQODLTWMAIG 61
D 1 SSKYPRSVRCCLPMALTLEAALLLFYFTTHDASLEDDKGLVASYQVQODLTWMAIG 60
QY 62 LGFTSSFRHSSSVAFNFMALGVQWAILLDGFLSOPPSGKVITLESILATMSAL 121
D 61 LGFTSSFRHSSSVAFNFMALGVQWAILLDGFLSOPPSGKVITLESILATMSAL 120
QY 122 SVLISDAVLGKYNLAQVYVAVLVEVTLGKLRVYISINFTDYHNMHIIYFAAYFGL 181
D 121 SVLISDAVLGKYNLAQVYVAVLVEVTLGKLRVYISINFTDYHNMHIIYFAAYFGL 180
QY 182 TVAMCLPKPLPEGETEDDQATTPSLSAMGALFLMFFSVSALLRSIERKNVNT 241
D 181 SVAMCLPKPLPEGETEDDQATTPSLSAMGALFLMFFSVSALLRSIERKNVNT 240
QY 242 YVAVAVSVVAIGSSSLAHPOGKISKTYGHSAYLPGCAVATGCHLIPSPWLPVLGLVA 301
D 241 YVAVAVSVVAIGSSSLAHPOGKISKTYGHSAYLPGCAVATGCHLIPSPWLPVLGLVA 300
QY 302 GLISVCAKTLPGCCNRYLGIPIHSSITGVNFSLLGLEIITIVLVLDVTGAGNGMIGF 361
D 301 GLISVCAKTLPGCCNRYLGIPIHSSITGVNFSLLGLEIITIVLVLDVTGAGNGMIGF 360
QY 362 QVLLSIGELSLAIVALTSGLTALNLKIRKAPHEAKYFDQVMPKPHLAVGF 417
D 361 QVLLSIGELSLAIVALTSGLTALNLKIRKAPHEAKYFDQVMPKPHLAVGF 416
RESULT 2
RHR_PANTR STANDARD; PRT; 416 AA.
AC Q28814;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE RH-LIKE PROTEIN IIR (RHESUS-LIKE PROTEIN IIR).
OS Pan troglodytes (Chimpanzee).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Pan.
OX NCBI_TaxID=9598;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-BONE MARROW.
RX MEDLINE=95085595; PubMed=7993375;
RA Salvignol I., Blancher A., Calvas P., Clayton J., Socha W.W.,
RA Colin Y., Ruffie J.;
RT "Molecular genetics of chimpanzee Rh-related genes: their
relationship with the R-C-E-F blood group system, the chimpanzee
counterpart of the human rhesus system.";
RL Biochem. Genet. 32:201-221(1994).
CC -1- FUNCTION: MAY BE PART OF AN OLIGOMERIC COMPLEX WHICH IS LIKELY TO
HAVE A TRANSPORT OR CHANNEL FUNCTION IN THE ERYTHROCYTE MEMBRANE.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -1- SIMILARITY: BELONGS TO THE RH FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -

Db 361 QVLSIGELSLAIVALTSLTGLTLNLIKIMKAPHAKEYDDQVEMKPPHLAVGF 416

RESULT 4

RH1A_PANTR STANDARD: PRT: 416 AA.

ID RH1A_PANTR

AC Q28813;

DT 01-NOV-1997 (Rel. 35, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)

DT 01-NOV-1997 (Rel. 35, Last annotation update)

DE RH-LIKE PROTEIN 1A (RHESUS-LIKE PROTEIN 1A).

OS Pan troglodytes (Chimpanzee).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Pan.

OC NCBI_TaxID=9598;

OX [1]

RP SEQUENCE FROM N.A.

RC TISSUE-BONE MARROW;

RX MEDLINE=95085595; PubMed-7993375;

RA Salvignol I., Blancher A., Calvas P., Clayton J., Socha W.W., Collin Y., Ruffie J.;

RT "Molecular genetics of chimpanzee Rh-related genes: their relationship with the R-C-E-F blood group system, the chimpanzee counterpart of the human rhesus system.";

RL Biochem. Genet. 32:201-221(1994).

CC -1- FUNCTION: MAY BE PART OF AN OLIGOMERIC COMPLEX WHICH IS LIKELY TO HAVE A TRANSPORT OR CHANNEL FUNCTION IN THE ERYTHROCYTE MEMBRANE.

CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.

CC -1- SIMILARITY: BELONGS TO THE RH FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

DR EMBL: L37049; AAA65623.1; -

DR InterPro: IPR001905; -

DR InterPro: IPR002229; -

DR Pfam: PF00909; Ammonium_transp; 1.

DR PRINTS: PR00342; RHESUSRHD.

DR Erythrocyte; Transmembrane.

KW INIT MET 0 0 BY SIMILARITY.

FT TRANSMEM 11 31 POTENTIAL.

FT TRANSMEM 43 63 POTENTIAL.

FT TRANSMEM 76 96 POTENTIAL.

FT TRANSMEM 124 144 POTENTIAL.

FT TRANSMEM 171 191 POTENTIAL.

FT TRANSMEM 202 222 POTENTIAL.

FT TRANSMEM 237 257 POTENTIAL.

FT TRANSMEM 264 284 POTENTIAL.

FT TRANSMEM 286 306 POTENTIAL.

FT TRANSMEM 330 350 POTENTIAL.

FT TRANSMEM 357 377 POTENTIAL.

SO SEQUENCE 416 AA; 45461 MW; B0B566734DB5E14D CRC64;

Query Match 86.4%; Score 1835; DB 1; Length 416;

Best Local Similarity 88.7%; Pred. No. 1,8e-126;

Matches 369; Conservative 11; Mismatches 36; Indels 0; Gaps 0;

QY 2 SCYPRSVGCLPLCALTLLEAAILLFFFTTHDASLEQKGLVASVQVQODLTVAATG 61

DB 1 SSRYPRSVRCLEPLCALTLLEAAILLFFFTQYDASLEQKGLVASVQVQODLTVAATG 60

QY 62 LGFLTSFRSHSSVAFNFLMALGYQMAILLDGFSLQPPSGKVVYTLFSLATMSAL 121

DB 61 FGLTSSFRSHSSVAFLFMAALGYQMAILLDGFSLQPPSGKVVYTLFSLATMSAL 120

QY 122 SVLISDAVLGKVNLAQLVVVLVEVTDGLNLRMVISNIFNTDYMNNMHIYFAAYFGL 181

Db 121 SVLISDAVLGKVNLAQLVVVLVEVTDGLNLRMVISNIFNTDYMNNMHIYFAAYFGL 180

QY 182 TVAMCLPKPLPECTEDNDQATIPSLAMIGALFLMMFPPSVNSALLRSPTEKKNVPM 241

DB 181 SVAMCLPKPLPECTEDNDQATIPSLAMIGALFLMMFPPSVNSALLRSPTEKKNVPM 240

QY 242 YVAAVSVYTAISGSSLAHPQCKISKTYGSAVLPQGVAVDTSCHLIPSPMLPVLGLVA 301

DB 241 YVAAVSVYTAISGSSLAHPQCKISKTYGSAVLPQGVAVDTSCHLIPSPMLPVLGLVA 300

QY 302 GLISVYGAKYLPCCCKNRVLGIPHSIMGYNFSLGLLEIYIVLVLDYVAGANGMIGF 361

DB 301 GLISIGAGKALPCCCKNRVLGIPHSIMGYNFSLGLLEIYIVLVLDYVAGANGMIGF 360

QY 362 QVLSIGELSLAIVALTSLTGLTLNLIKIMKAPHAKEYDDQVEMKPPHLAVGF 417

DB 361 QVLRIGERSLATYITLALTSLTGLTLNLIKIMKAPHAKEYDDQVEMKPPHLAVGF 416

RESULT 5

RHCE_HUMAN STANDARD: PRT: 416 AA.

ID RHCE_HUMAN

AC P18577; Q02163; Q02164; Q02165; Q16160;

DT 01-NOV-1990 (Rel. 16, Created)

DT 01-NOV-1990 (Rel. 16, Last sequence update)

DT 01-OCT-2000 (Rel. 40, Last annotation update)

DE BLOOD GROUP RH(C)E POLYPEPTIDE (RHESUS C/E ANTIGENS) (RH30A) (RH1XB) (RH POLYPEPTIDE 1) (RHPI).

OS RHCE OR RHC OR RHE.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.

OC NCBI_TaxID=9606;

OX [1]

RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.

RC TISSUE-BONE MARROW;

RX MEDLINE=90349591; PubMed-1696722;

RA Cherif-Zahar B., Bloy C., Le van Kim C., Blanchard D., Bailly P., Hernand P., Salmon C., Cartton J.-P., Collin Y.;

RT "Molecular cloning and protein structure of a human blood group Rh polypeptide.";

RL Proc. Natl. Acad. Sci. U.S.A. 87:6243-6247(1990).

RN [2]

RP SEQUENCE FROM N.A.

RX MEDLINE=91058522; PubMed-2123099;

RA Avent N.D., Ridgwell K., Tanner M.J.A., Anstee D.J.;

RT "cDNA cloning of a 30 kDa erythrocyte membrane protein associated with Rh (Rhesus)-blood-group-antigen expression.";

RL Biochem. J. 271:821-825(1990).

RN [3]

RP SEQUENCE FROM N.A.

RX MEDLINE=93216282; PubMed-7916743;

RA Kajli E., Umenishi F., Iwamoto S., Ikemoto S.;

RT "Isolation of a new cDNA clone encoding an Rh polypeptide associated with the Rh blood group system.";

RL Hum. Genet. 91:157-162(1993).

RN [4]

RP SEQUENCE FROM N.A. (FORMS RHIV AND RHVIII).

RC TISSUE-BONE MARROW;

RX MEDLINE=92360855; PubMed-1379850;

RA le van Kim C., Cherif-Zahar B., Raynal V., Mouro I., Lopez M., Cartton J.-P., Collin Y.;

RT "Multiple Rh messenger RNA isoforms are produced by alternative splicing.";

RL Blood 80:1074-1078(1992).

RN [5]

RP SEQUENCE OF 1-32.

RX MEDLINE=89134163; PubMed-3146980;

RA Avent N.D., Ridgwell K., Mawby W.J., Tanner M.J.A., Anstee D.J., Kumpel B.;

RT "Protein-sequence studies on Rh-related polypeptides suggest the presence of at least two groups of proteins which associate in the human red-cell membrane.";


```

ID  RULE_PANTR  STANDARD;  PRT;  416 AA.
AC  Q28426;
DT  01-NOV-1997 (Rel. 35, Created)
DT  01-NOV-1997 (Rel. 35, Last sequence update)
DE  RH-LIKE PROTEIN IIF (RHESUS-LIKE PROTEIN IIF).
OS  Pan troglodytes (Chimpanzee).
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Pan.
OX  NCBI_TaxID=9598;
RN  [1]
RP  SEQUENCE FROM N.A.
RX  MEDLINE=95085595; PubMed=7993375;
RA  Salvignol I., Blancher A., Calvas P., Clayton J., Socha W.W.,
RA  Colin Y., Ruffie J.;
RT  "Molecular genetics of chimpanzee Rh-related genes: their
RT  relationship with the R-C-E-F blood group system, the chimpanzee
RT  counterpart of the human rhesus system.";
RL  Biochem. Genet. 32:201-221(1994).
CC  -1- FUNCTION: MAY BE PART OF AN OLIGOMERIC COMPLEX WHICH IS LIKELY TO
CC  HAVE A TRANSPORT OR CHANNEL FUNCTION IN THE ERYTHROCYTE MEMBRANE.
CC  -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC  -1- SIMILARITY: BELONGS TO THE RH FAMILY.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL; L37048; AAA65622.1; -
DR  InterPro; IPR001905; -
DR  Pfam; PF00909; Ammonium_transp. 1.
DR  PRINTS; PR00342; RHESUSRD.
KW  Erythrocyte; Transmembrane.
FT  INIT_MET 0
FT  TRANSMEM 11 31
FT  TRANSMEM 43 63
FT  TRANSMEM 76 96
FT  TRANSMEM 124 144
FT  TRANSMEM 171 191
FT  TRANSMEM 202 222
FT  TRANSMEM 237 257
FT  TRANSMEM 264 284
FT  TRANSMEM 286 306
FT  TRANSMEM 330 350
FT  TRANSMEM 357 377
SQ  SEQUENCE 416 AA; 45136 MW; 325ED916268BAF2F CRC64;

Query Match 85.0%; Score 1806; DB 1; Length 416;
Best Local Similarity 87.5%; Pred. No. 2,3e-124;
Matches 364; Conservative 11; Mismatches 41; Indels 0; Gaps 0;

```

```

OY 242 YVAVAVSVTAISGSSLAHPDGKISKTYGSHAVLPGECVAVDTSCHLIPSPWLPVIGLVA 301
DB 241 YVLAASVTVTAISGSSLAHPDGKISKMTYVHSAVLGAVGTSCHLIPSPWLAIVGLVA 300
OY 302 GLISVIGAKLPCCNRVLCIPHSIMGVNPSLLGLLEITIVLVLPVGVGNMIGF 361
DB 301 GLISIGAKGPCCNVLCIPDSVWHYVNSLGLLEITIVLVLRHTVWNGNMIGF 360
OY 362 QVLLSIGELSLAIVALTSLGALLNLKIRKAPAEKEDDOVPMKPFHLAVGF 417
DB 361 QVLLSGELSLAIVALTSLGALLNLKIRKAPAEKEDDOVPMKPFHLAVGF 416

RESULT 7
RHLC_GORGO
ID  RHLC_GORGO  STANDARD;  PRT;  416 AA.
AC  Q28426;
DT  01-NOV-1997 (Rel. 35, Created)
DT  01-NOV-1997 (Rel. 35, Last sequence update)
DE  RH-LIKE PROTEIN IC (RHESUS-LIKE PROTEIN IC).
OS  Gorilla gorilla gorilla (Lowland gorilla).
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Gorilla.
OX  NCBI_TaxID=9595;
RN  [1]
RP  SEQUENCE FROM N.A.
RX  MEDLINE=95085595; PubMed=7993375;
RA  Salvignol I., Blancher A., Calvas P., Clayton J., Socha W.W.,
RA  Colin Y., Ruffie J.;
RT  "Molecular genetics of chimpanzee Rh-related genes: their
RT  relationship with the R-C-E-F blood group system, the chimpanzee
RT  counterpart of the human rhesus system.";
RL  Biochem. Genet. 32:201-221(1994).
CC  -1- FUNCTION: MAY BE PART OF AN OLIGOMERIC COMPLEX WHICH IS LIKELY TO
CC  HAVE A TRANSPORT OR CHANNEL FUNCTION IN THE ERYTHROCYTE MEMBRANE.
CC  -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC  -1- SIMILARITY: BELONGS TO THE RH FAMILY.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL; L37052; AAA65626.1; -
DR  InterPro; IPR001905; -
DR  Pfam; PF00909; Ammonium_transp. 1.
DR  PRINTS; PR00342; RHESUSRD.
KW  Erythrocyte; Transmembrane.
FT  INIT_MET 0
FT  TRANSMEM 11 31
FT  TRANSMEM 43 63
FT  TRANSMEM 76 96
FT  TRANSMEM 124 144
FT  TRANSMEM 171 191
FT  TRANSMEM 202 222
FT  TRANSMEM 237 257
FT  TRANSMEM 264 284
FT  TRANSMEM 286 306
FT  TRANSMEM 330 350
FT  TRANSMEM 357 377
SQ  SEQUENCE 416 AA; 45129 MW; 0F3B623F9908E087 CRC64;

Query Match 84.6%; Score 1798; DB 1; Length 416;
Best Local Similarity 86.5%; Pred. No. 8.7e-124;
Matches 360; Conservative 15; Mismatches 41; Indels 0; Gaps 0;

```



```

RL Biochem. Genet. 32:201-221(1994).
CC -1- FUNCTION: MAY BE PART OF AN OLIGOMERIC COMPLEX WHICH IS LIKELY TO
CC HAVE A TRANSPORT OR CHANNEL FUNCTION IN THE ERYTHROCYTE MEMBRANE.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -1- SIMILARITY: BELONGS TO THE RH FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: L37054; AAA65628.1;
CC InterPro: IPR001905;
CC Pfam: PF00909; Ammonium_transp; 1.
CC PRINTS: PR00342; RHESUSRD.
CC Erythrocyte; Transmembrane.
CC KW INIT_MET 0 BY SIMILARITY.
CC FT TRANSMEM 11 31 POTENTIAL.
CC FT TRANSMEM 43 63 POTENTIAL.
CC FT TRANSMEM 76 96 POTENTIAL.
CC FT TRANSMEM 124 144 POTENTIAL.
CC FT TRANSMEM 171 191 POTENTIAL.
CC FT TRANSMEM 202 222 POTENTIAL.
CC FT TRANSMEM 237 257 POTENTIAL.
CC FT TRANSMEM 264 284 POTENTIAL.
CC FT TRANSMEM 286 306 POTENTIAL.
CC FT TRANSMEM 330 350 POTENTIAL.
CC FT TRANSMEM 357 377 POTENTIAL.
CC SQ SEQUENCE 416 AA; 45877 MW; 17822A558D299A9 CRC64;

Query Match 70.8%; Score 1504; DB 1; Length 416;
Best Local Similarity 74.3%; Pred. No. 2e-102;
Matches 309; Conservative 31; Mismatches 76; Indels 0; Gaps 0;

QY 2 SCYTPRSVGRCLPLCALTEALILFEFTHYDASLEDDKGLVASYOGODLTVAAG 61
DB 1 SKYTPRSVGRCLPLMALTEALILFEFTTYDASLEDDKGLVASYOGODLTVAAG 60
QY 62 LGLTSSFRHSHSSVAFLNFMALGVOMAILLDGFLSOPSGKVVITLPSIATMSAL 121
DB 61 LGLTSSFRHSHSSVAFLNFMALGVOMAILLDGFLSOPSGKVVITLPSIATMSAL 120
QY 122 SVLISDAVYLGKYNLAQLVYVLEVTGLGNLRMVISNIFNDYHNNMMHIYFAAYEGL 181
DB 121 SMLISMAVYLGKYNLAQLVYVLEVTGLGNLRMVISNIFNDYHNNMMHIYFAAYEGL 180
QY 182 TVAMCLPKPLPEGETEDNDCAFIPSLSAMGALFLMMPFSPVSALLRSPRIERKNVFN 241
DB 181 TVAMCLPKPLPEGETEDNDCAFIPSLSAMGALFLMMPFSPVSALLRSPRIERKNVFN 240
QY 242 YVAVAVSVAITAISSSLAHPOGKISKTYGHSAYLPBGAVADTSCHLIPSPWLPYVIGLVA 301
DB 241 YVAVAVSVAITAISSSLAHPOGKISKTYGHSAYLPBGAVADTSCHLIPSPWLPYVIGLVA 300
QY 302 GLISVGAATLPGCCNRYLCIPIHSSINGVYFSLGLLEIYIVLVLDYVAGNGNGIGF 361
DB 301 GLISVGAATLPGCCNRYLCIPIHSSINGVYFSLGLLEIYIVLVLDYVAGNGNGIGF 360
QY 362 QVLTISGELSLAVITLSTGLTALLNLKIRKAPHAKEFDVDFPKPHLAVGF 417
DB 361 QVLTISGELSLAVITLSTGLTALLNLKIRKAPHAKEFDVDFPKPHLAVGF 416

RESULT 10
RHL_HYLP1 STANDARD; PRT; 353 AA.
AC Q28446;
DT 01-NOV-1997 (Rel. 35, Created)

```

```

DT 01-NOV-1997 (Rel. 35, Last sequence update)
DE 01-NOV-1997 (Rel. 35, Last annotation update)
DE RH-LIKE PROTEIN (RHESUS-LIKE PROTEIN).
OS Hylobates pileatus (Pileated gibbon).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hylobatidae; Hylobates.
ON NCBI_TaxID=9589;
RX [1]
RC TISSUE-BONE marrow;
RX MEDLINE=95085595; PubMed=7993375;
RA Salviagnol I., Blancher A., Calvas P., Clayton J., Socia W.W.,
RA Collin Y., Ruffie J.;
RT "Molecular genetics of chimpanzee Rh-related genes: their
RT relationship with the R-C-E-F blood group system, the chimpanzee
RT counterpart of the human rhesus system.";
RT Biochem. Genet. 32:201-221(1994).
CC -1- FUNCTION: MAY BE PART OF AN OLIGOMERIC COMPLEX WHICH IS LIKELY TO
CC HAVE A TRANSPORT OR CHANNEL FUNCTION IN THE ERYTHROCYTE MEMBRANE.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -1- SIMILARITY: BELONGS TO THE RH FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: L37051; AAA65625.1;
CC InterPro: IPR001905;
CC Pfam: PF00909; Ammonium_transp; 1.
CC Erythrocyte; Transmembrane.
CC KW INIT_MET 0 BY SIMILARITY.
CC FT TRANSMEM 10 30 POTENTIAL.
CC FT TRANSMEM 44 64 POTENTIAL.
CC FT TRANSMEM 76 96 POTENTIAL.
CC FT TRANSMEM 124 144 POTENTIAL.
CC FT TRANSMEM 166 186 POTENTIAL.
CC FT TRANSMEM 208 228 POTENTIAL.
CC FT TRANSMEM 237 257 POTENTIAL.
CC FT TRANSMEM 286 306 POTENTIAL.
CC SQ SEQUENCE 353 AA; 38417 MW; 3CFD25B38033AB6A CRC64;

Query Match 60.7%; Score 1289; DB 1; Length 353;
Best Local Similarity 82.7%; Pred. No. 7.2e-87;
Matches 258; Conservative 20; Mismatches 34; Indels 0; Gaps 0;

QY 2 SCYTPRSVGRCLPLCALTEALILFEFTHYDASLEDDKGLVASYOGODLTVAAG 61
DB 1 SKYTPRSVGRCLPLMALTEALILFEFTTYDASLEDDKGLVASYOGODLTVAAG 60
QY 62 LGLTSSFRHSHSSVAFLNFMALGVOMAILLDGFLSOPSGKVVITLPSIATMSAL 121
DB 61 LGLTSSFRHSHSSVAFLNFMALGVOMAILLDGFLSOPSGKVVITLPSIATMSAL 120
QY 122 SVLISDAVYLGKYNLAQLVYVLEVTGLGNLRMVISNIFNDYHNNMMHIYFAAYEGL 181
DB 121 SMLISMAVYLGKYNLAQLVYVLEVTGLGNLRMVISNIFNDYHNNMMHIYFAAYEGL 180
QY 182 TVAMCLPKPLPEGETEDNDCAFIPSLSAMGALFLMMPFSPVSALLRSPRIERKNVFN 241
DB 181 TVAMCLPKPLPEGETEDNDCAFIPSLSAMGALFLMMPFSPVSALLRSPRIERKNVFN 240
QY 242 YVAVAVSVAITAISSSLAHPOGKISKTYGHSAYLPBGAVADTSCHLIPSPWLPYVIGLVA 301
DB 241 YVAVAVSVAITAISSSLAHPOGKISKTYGHSAYLPBGAVADTSCHLIPSPWLPYVIGLVA 300
QY 302 GLISVGAATLPG 313
DB 301 GLISVGAATLPG 312

```


OS Escherichia coli.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Escherichia.
 OX NCBI_TaxID=562;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=88273145; PubMed=2839489;
 RA Karpel R., Olami Y., Taglicht D., Schuldiner S., Padan E.;
 RT "Sequencing of the gene ant which affects the Na⁺/H⁺ antiporter
 activity in Escherichia coli.";
 RL J. Biol. Chem. 263:10408-10414(1988).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-K12;
 RX MEDLINE=92334977; PubMed=1630901;
 RA Yura T., Mori H., Nagai H., Nagata T., Ishihama A., Fujita N.,
 Isoono K., Mizobuchi K., Nakata A.;
 RT "Systematic sequencing of the Escherichia coli genome: analysis of
 the 0-2.4 min region.";
 RL Nucleic Acids Res. 20:3305-3308(1992).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-K12/MG1655;
 RX MEDLINE=97426617; PubMed=9278503;
 RA Blatter F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
 Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
 Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
 Mau B., Shao Y.;
 RT "The complete genome sequence of Escherichia coli K-12.";
 RL Science 277:1233-1238(1997).
 RN [4]
 RP SEQUENCE OF 1-28 FROM N.A.
 RX MEDLINE=92042080; PubMed=1657980;
 RA Karpel R., Alon T., Glaser G., Schuldiner S., Padan E.;
 RT "Expression of a sodium proton antiporter (Nhaa) in Escherichia coli
 is induced by Na⁺ and Li⁺ ions.";
 RL J. Biol. Chem. 266:21753-21759(1991).
 RN [5]
 RP FUNCTION, AND PARTIAL SEQUENCE.
 RX MEDLINE=91250446; PubMed=1645730;
 RA Taglicht D., Padan E., Schuldiner S.;
 RT "Overproduction and purification of a functional Na⁺/H⁺ antiporter
 coded by nhaa (ant) from Escherichia coli.";
 RL J. Biol. Chem. 266:11289-11294(1991).
 RN [6]
 RP MUTAGENESIS OF HISTIDINES.
 RX MEDLINE=93165668; PubMed=8381959;
 RA Gerschman Y., Olami Y., Rimon A., Taglicht D., Schuldiner S., Padan E.;
 RT "Histidine-226 is part of the pH sensor of Nhaa, a Na⁺/H⁺ antiporter
 in Escherichia coli.";
 RL Proc. Natl. Acad. Sci. U.S.A. 90:1212-1216(1993).
 CC -1- FUNCTION: THIS IS A Na⁺/H⁺ ANTI-PORTER. IT CAN MEDIATE SODIUM
 UPTAKE WHEN A TRANSMEMBRANE PH GRADIENT IS APPLIED. IT PROBABLY
 ACTS IN THE REGULATION OF INTERNAL PH AT THE ALKALINE PH RANGE. IT
 CATALYZES THE EXCHANGE OF 2H⁺ PER NA⁺. ITS ACTIVITY IS HIGHLY
 DEPENDENT ON THE PH.
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. INNER MEMBRANE.
 CC -1- INDUCTION: TRANSCRIPTION STIMULATED BY HIGH NA⁺ CONCENTRATIONS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: J03879; AAA23448.1; ALT_INIT.
 CC EMBL: D10483; BAA01296.1; ALT_INIT.
 CC EMBL: AEO00112; AAC73130.1; -.
 CC EMBL: S67239; AAB20348.1; -.
 CC EMBL: A28800; A28800.
 CC PIR: S40541; S40541.

DR Ecogene; EGI0652; nhaa.
 KW Transmembrane; Inner membrane; Transport; Sodium transport; Symport.
 FT TRANSMEM 59 79 POTENTIAL.
 FT TRANSMEM 95 115 POTENTIAL.
 FT TRANSMEM 125 145 POTENTIAL.
 FT TRANSMEM 154 174 POTENTIAL.
 FT TRANSMEM 179 199 POTENTIAL.
 FT TRANSMEM 205 221 POTENTIAL.
 FT TRANSMEM 254 274 POTENTIAL.
 FT TRANSMEM 283 299 POTENTIAL.
 FT TRANSMEM 328 348 POTENTIAL.
 FT TRANSMEM 360 380 POTENTIAL.
 FT SITE 225 225 IMPORTANT FOR PH SENSOR.
 FT CONFLICT 56 57 NK -> KQ (IN REF. 1).
 SQ SEQUENCE 388 AA; 41355 MW; B508B1D2E5EE9130 CRC64;
 Query Match 5.6%; Score 119.5; DB 1; Length 388;
 Best Local Similarity 20.0%; Pred. No. 0.081;
 Matches 82; Conservative 69; Mismatches 143; Indels 117; Gaps 16;
 QY 18 LLEALILFFPTTHDASLEQKGLVASYQVGDLTVAATG-----LGFITSSF-- 69
 DB 61 LWINDALMAVFLVGLVEKRELMQSGSLAQAA-FPVIAAIGMIVPALTYLAFTNAD 119
 QY 70 -RRHSW-----SSVAFNLEFLA-LGVQMAILLDFLSQFP-----SGKVITLF--SI 113
 DB 120 PIRREGAIPRAADIANALGVALLLGSRVPLAKIFLMALAIIDDGAILIILFTINDL 179
 QY 114 WLATMSALSVLISDAVLAQVLAQVVMVLEVTDLGNRMVINSINFTDYIMNMHIY 173
 DB 180 SMASLGVAAVAIAVLAVLNLGARGGVYILGV-----VIMTAVLKSQVH----- 225
 QY 174 VPAVREGTLVANKLPKPEGTEDNDQRTIPISLAMGLFLMMFPPSNALSLSPIE 233
 DB 226 -ATLGVIVGFPIPLKEKRG-----NSPAK 249
 QY 234 RKNVENVTYAVAVSVYVTAISGSSLAHPQKISKTYGHSVAVLDEGVAVDTSCHLIPSPWL 293
 DB 250 RLEHVLHPWVAVYLLPLFA-----FANAGVSLOQVTLDOGLTSLP---- 289
 QY 294 PVLGVAGLISVYAKYLPDCCNRVLGPRHSSIMGYNFSLCLLEIYIVLLVDTV- 352
 DB 290 ---LGIITAGLL-----IGKPLGISLPCWLAIRLKLAIHPGTYQQQIMVVGIIIC 335
 QY 353 GAGNGMIGPVLLSIGLSLAIV-----IALTSGLTALILN---LKIRKAP 396
 DB 336 GIGFTMSIFLASLAFGSDPELINMAKLGITLVGSISSAVIGYSWLRVLRP 386
 RESULT 13
 ID YJ78_YEAST STANDARD; PRT; 881 AA.
 AC P39535;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE HYPOHETICAL 97.7 KDA MEMBRANE PROTEIN IN PRP21-UBP12 INTERGENIC
 DE REGION.
 GN YJL198W OR J0336.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=5288C;
 RX MEDLINE=95274326; PubMed=7754713;
 RA Purnelle B., Coster F., Goffeau A.;
 RT "The sequence of a 36 kb segment on the left arm of yeast chromosome
 X identifies 24 open reading frames including NUC1, PRP21 (SPP1),
 RT CDC6, CRY2, the gene for S24, a homologue to the acetylase gene ACO1
 and two homologues to chromosome III genes.";

RL Yeast 10:1235-1249(1994).
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).
 CC -1- SIMILARITY: BELONGS TO THE NAD(P)/PHO87 FAMILY OF TRANSPORTERS.
 CC PHO87 SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X77688; CAA54759.1; -
 DR EMBL: 249473; CAA89493.1; -
 DR PIR: S46633; S46633.
 DR SCD: S0003734; YUL198W.
 KW Hypothetical protein; Transmembrane; Transport.
 KM Hypothetical protein; Transmembrane; Transport.
 FT TRANSMEM 417 437 POTENTIAL.
 FT TRANSMEM 456 476 POTENTIAL.
 FT TRANSMEM 493 513 POTENTIAL.
 FT TRANSMEM 514 534 POTENTIAL.
 FT TRANSMEM 539 559 POTENTIAL.
 FT TRANSMEM 581 601 POTENTIAL.
 FT TRANSMEM 663 683 POTENTIAL.
 FT TRANSMEM 691 711 POTENTIAL.
 FT TRANSMEM 718 738 POTENTIAL.
 FT TRANSMEM 758 778 POTENTIAL.
 FT TRANSMEM 805 825 POTENTIAL.
 FT TRANSMEM 854 874 POTENTIAL.
 SQ SEQUENCE 881 AA; 97688 MW; BB4466BCAD176714 CRC64;

Query Match 5.6%; Score 119; DB 1; Length 881;
 Best Local Similarity 20.5%; Pred. No. 0.21; Mismatches 130; Indels 116; Gaps 17;

Matches 80; Conservative 65; Mismatches 130; Indels 116; Gaps 17;
 QY 86 LGVOMA-----ILDGFE-----LSQFPGSKV-----VITLFSIW 114
 DB 494 LAAMWSSTIMILLAGFTLGEVLAQYNAKVLASWLAFAQCKPRNVLLMAMCVFFFLSMW 553
 QY 115 LA-----TMSALSVLSVDAVIGKYNLAQVVMVLEVTGDLNLRMVISINIFNTDYM 167
 DB 554 ISNVAALVLYSTLSPL--DAMDADSPFAQALVGLVALAANIGMSSPISPNONI--I 608
 QY 168 NMHIYFAAYFGLTVMKCPKPLPEGTEDDDQATIPSLAMGALFLMW-----TPPS 222
 DB 609 SMSYLPKPYGIGWGQFFAVALP-----SGILAMLLWLILFTTKMN 649
 QY 223 VNSALRSPIERKNAVENTYAAVAVVTAISGSSLAHPQGRISKTYGHSVLP-----EG 278
 DB 650 KTKLEKRPKIKTKTV-KQYIIIVYATILLMCVESQIEBAPSS-QQAIIRIVLFPFG 707
 QY 279 VAVDTSCHLIPSPWLPIVLGT-----VAGLISVYGAQYLPG-----CCN 317
 DB 708 TGLSTODLNAFPIVYIILAMGIALGKAVSSGSLSTI-AKALQKRIENDGVAIICIF 766
 QY 318 RVIGIPHSSTMGVNF--LLGLLEF-----IYIVLVLDYVAGNGMIGFO 362
 DB 767 GILMLVGVTFVSHVSAIIIIPIVQEVGDKLGNPKAAPILVGCALLSSCGMGLASSGFP 826
 QY 363 VLSTI-----GELSLAVIALTSGLLTALL 387
 DB 827 NVTAISKVDRKGRKSLVMTFLTRGVPASIL 857
 RESULT 14
 Y092_HAEIN STANDARD; PRT; 419 AA.
 AC 057493;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)

DE HYPOTHETICAL PROTEIN HI0092.
 GN HI0092.
 OS Haemophilus influenzae.
 OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
 CC Haemophilus.
 CC NCBI_TaxID=727;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC SPRAIN-RD / KW20 / ATCC 51907;
 RX MEDLINE=95350630; Pubmed=7542800;
 RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
 RA Kierulff A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
 RA Mckenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
 RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
 RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
 RA Ufferback T.R., Hanna M.C., Nguyen D.T., Sauek D.M., Brandon R.C.,
 RA Fine L.D., Frichman J.L., Fuhmann J.L., Geoghagen N.S.M.,
 RA Gnehm C.U., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
 RA Venter J.C.,
 RA "Whole-genome random sequencing and assembly of Haemophilus
 RA influenzae Rd.";
 RT Science 269:496-512(1995).
 RL -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: U32694; AAC21770.1; -
 DR TIGR: HI0092; -
 KW Hypothetical protein; Transmembrane.
 KM Hypothetical protein; Transmembrane.
 FT TRANSMEM 1 21 POTENTIAL.
 FT TRANSMEM 24 44 POTENTIAL.
 FT TRANSMEM 66 86 POTENTIAL.
 FT TRANSMEM 101 121 POTENTIAL.
 FT TRANSMEM 174 194 POTENTIAL.
 FT TRANSMEM 216 236 POTENTIAL.
 FT TRANSMEM 242 262 POTENTIAL.
 FT TRANSMEM 280 300 POTENTIAL.
 FT TRANSMEM 311 331 POTENTIAL.
 FT TRANSMEM 349 369 POTENTIAL.
 FT TRANSMEM 396 416 POTENTIAL.
 SQ SEQUENCE 419 AA; 42781 MW; 6DCA54BB0A55FF73 CRC64;

Query Match 5.3%; Score 112.5; DB 1; Length 419;
 Best Local Similarity 19.8%; Pred. No. 0.28; Mismatches 145; Indels 113; Gaps 17;
 Matches 81; Conservative 71; Mismatches 145; Indels 113; Gaps 17;

QY 43 GLVASYGVGDVYMAAIGLGTLSFRHRSSVSAFNLMLALGVOMAILD-----GFL 98
 DB 38 GLIGCALDSQVTSIMIGAGCITTAVMK-----ILAGVLAGVLDIESGAANSI 85
 QY 99 SQFSGKVVIT--LESLATMSALSVISYDAVIGKYNLAQVVMVLEVTGDLNLRMV 156
 DB 86 TETITNKLGETRALLALALAMILTAAGVFDVAIVTS--PALASRRSDLSKAAIL 142
 QY 157 IS-----NIFNTD-----YHM--NMHIYFAAYFGLTVMKCPKPL-DEGT 195
 DB 143 LAMIGGKAGKAGNIMSPNPALAAADTFHPLPLSVMMAGIIPALFOLLILYFLAKLINKGS 202
 QY 196 EDNDQFA-----TIPS--LSAMGALFLMFP-PSVNSALRSPIERKNAVENTYAAV 245
 DB 203 KVTDEKIVYLETQNLPSFLALVAPLVAIILALARP-----LFDI 242
 QY 246 AVSVVTAISGSSL-AHPQGRISKTYGHSNAVLPQGVAVDTSCHLIPSPWLPIVLGVL 303
 DB 243 KVDPLIALPLGLGLTGACMGKLRINISY-----AINGLSKMTPVAIMLGTGALAGI 294

OY	363	ISV:CAKULPGCCNRVLGIPHSISXNGVFNLSGLLEIYIYLVLDYVAGNGMIGROY	363
DB	295	IANSLSKEY-----LIQIENHSGLEST-----ILAPISG-----V	324
OY	364	LSISGELSLAIVIALTSGILATALLNLKIRKAPHEAK-YFDQVPEWKEPPH	412
DB	325	LSMSLATPSTAGTAVASNVFSSTLEICVSLAGAAAMHAGATVPDHPH	374
RESULT	15		
ID	GNTT_ECOLI	STANDARD:	PRT: 437 AA.
AC	P39835:		
DT	01-FEB-1995 (Rel. 31, Created)		
DT	01-NOV-1995 (Rel. 32, Last sequence update)		
DT	01-NOV-1997 (Rel. 35, Last annotation update)		
DE	HIGH-AFFINITY GLUCONATE TRANSPORTER (GLUCONATE PERMEASE) (GNT-I SYSTEM).		
GN	GNTT OR USGA OR GNTM.		
OS	Escherichia coli.		
OC	Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;		
OX	NCBI_TaxID=562;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN-K12 / MG1655;		
RX	MEDLINE=97426617; PubMed=9278503;		
RA	Blatter F.R., Plunkett G. II, Bloch C.A., Perna N.T., Burland V.,		
RA	Riley M., Collado-Valdes J., Glasner J.D., Rode C.K., Mayhew G.F.,		
RA	Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,		
RA	Mau B., Shao Y.;		
RT	"The complete genome sequence of Escherichia coli K-12.";		
RL	Science 277:1453-1474(1997).		
RN	[2]		
RP	PRELIMINARY SEQUENCE OF 381-437 FROM N.A.		
RX	MEDLINE=89013888; PubMed=2845225;		
RA	Pugsley A.P., Dubreuil C.;		
RT	"Molecular characterization of malQ, the structural gene for the		
RT	Escherichia coli enzyme amyloamylase.";		
RL	Mol. Microbiol. 2:473-479(1988).		
RN	[3]		
RP	IDENTIFICATION.		
RX	MEDLINE=95075659; PubMed=7984428;		
RA	Borodovsky M., Rudd K.E., Koonin E.V.;		
RT	"Intrinsic and extrinsic approaches for detecting genes in a		
RT	bacterial genome.";		
RL	Nucleic Acids Res. 22:4756-4767(1994).		
RN	[4]		
RP	CHARACTERIZATION.		
RX	MEDLINE=87310381; PubMed=3040894;		
RA	Isturiz T., Palmero E., Vitelli-Flores J.;		
RT	"Mutations affecting gluconate catabolism in Escherichia coli.		
RT	Genetic mapping of the locus for the thermosensitive gluconokinase.";		
RL	J. Gen. Microbiol. 132:3209-3219(1986).		
RN	[5]		
RP	CHARACTERIZATION.		
RX	MEDLINE=97197521; PubMed=9045817;		
RA	Porco A., Peeksh N., Bausch C., Tong S., Isturiz T., Conway T.;		
RT	"Molecular genetic characterization of the Escherichia coli gntI gene		
RT	of gntI, the main system for gluconate metabolism.";		
RL	J. Bacteriol. 179:1584-1590(1997).		
CC	-1- FUNCTION: PART OF THE GLUCONATE UTILIZATION SYSTEM GNT-I; HIGH-		
CC	AFFINITY INTAKE OF GLUCONATE.		
CC	-1- PATHWAY: GLUCONATE UTILIZATION.		
CC	-1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. INNER MEMBRANE		
CC	(POTENTIAL).		
CC	-1- SIMILARITY: BELONGS TO THE GNTP FAMILY OF PERMEASES.		
CC	-----		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration		
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation		
CC	the European Bioinformatics Institute. There are no restrictions on its		
CC	use by non-profit institutions as long as its content is in no way		
CC	modified and this statement is not removed. Usage by and for commercial		

CC	entiles requires a license agreement (See http://www.lsb-sib.ch/announce/					
CC	or send an email to license@sib-sib.ch). -----					
DR	EMBL; U18997; AAN8213.1; "-					
DR	EMBL; AE000417; AAC76440.1; "-					
DR	EMBL; M32793; "- NOT_ANNOTATED_CDS.					
DR	Ecogene; EG12380; gntT.					
KM	Gliconate utilization; Sugar transport;					
KM	Inner membrane; Multigene family.					
FT	TRANSMEM	2	22	POTENTIAL.		
FT	TRANSMEM	23	43	POTENTIAL.		
FT	TRANSMEM	50	70	POTENTIAL.		
FT	TRANSMEM	107	127	POTENTIAL.		
FT	TRANSMEM	133	153	POTENTIAL.		
FT	TRANSMEM	173	193	POTENTIAL.		
FT	TRANSMEM	221	241	POTENTIAL.		
FT	TRANSMEM	257	277	POTENTIAL.		
FT	TRANSMEM	291	311	POTENTIAL.		
FT	TRANSMEM	326	346	POTENTIAL.		
FT	TRANSMEM	348	368	POTENTIAL.		
FT	TRANSMEM	369	389	POTENTIAL.		
FT	TRANSMEM	417	437	POTENTIAL.		
FT	CONFLICT	384	384	S -> A (IN REF. 2).		
FT	CONFLICT	419	420	ET -> VS (IN REF. 2).		
SO	SEQUENCE	437 AA;	45923 MW;	2AEBF78BD0459D66 CAC64;		

Query Match	5.28;	Score 111.5;	DB 1;	Length 437;
Best Local Similarity	20.5%;	Pred. NO. 0.35;		
Matches 93;	Conservative 68;	Mismatches 152;	Indels 141;	Gaps 20;

```

01 56 VMAALGJGFJSSRRRRSSWSS--VAFNLFEMALQVOMILLDGLSOPSGKVITLFSW 114
02 4 VIVAIGVILLLLLMIRRMNGFIALVVALAVALGMSGMRPDKYIGSKAG----- 53
03 115 LATMSALSVLLSDVALGKV---NLAOLVVMVLEVTDLGNLR----- 154
04 54 VADGSLALINGFAMGKMLADCGAGATATTLIAFGKKNQOMAVLTLGTFVGFALEY 113
05 155 MYSNLFENFDYIMNMNMHIVY---FAAYGELYAMKLEPRPREGSTDDNDQATIFS 206
06 114 EVGEVLMRLPFTTAAASNPRLDLVGVYMAAASVYNGFLRPHRGPAI---ATI-- 165
07 207 L$AMIGALFEMFPR$VSYNALLRSPRIERKNAVENTYUAAVSVVTAISGSLAHPOKIS 266
08 166 FNADMGKLLGLTILALPTVILAPV-----YA-----RYLKGIDRKIRPGLS 209
09 267 -KTYGSHAVLPREGVAVGDTSCHLIRPMLPILV--GLVAGLISVFGAKYLRCCCNRYLGI 322
10 210 AKTSESEEMRPFSGSVYVSTL-----VPVULMAMRAIAEMIRPKGAPLP--VAEFLTD 260
11 323 PHSSIMGCVNFSLL-----GL-----LDELIVYLLVLDLPVVGNGNMIGFOYL 364
12 261 P---VMATLILAVLIAMFTFGLNCRSMQDINDTLVSSIKIIMAMLLIIGGGAF--KOYL 315
13 365 L$IG-----ELSLAIVALTSGILTALLGN----- 389
14 316 VDSGVDKYIASMHMETNISPLLMAM$IAAIVRIATAG$ATVAATLTAGIAPLIATGVSP 375
15 390 ----LKITKAPHEAKYFDDOVFNKFP---HLANG 416
16 376 ELMAVIAVGSGVIFSHVNDP$FMLEK$EYFNILTG 409

```

Search completed: September 12, 2001, 07:42:42
Job time: 211 sec

THIS PAGE BLANK (USPTO)


```

Db 121 LSVLISYDAVLGKYNLAQLVVMVLEVTALGNLRMVISNIFNTDYHMMMHIIYFAAYFG 180
QY 181 LTVAMCPLPKPLPECTEDDDQATIPSLSAMGLFLMFPVSVALRSPIERKNVFN 240
   |||||||
Db 181 LSVAMCPLPKPLPECTEDDDQATIPSLSAMGLFLMFPVSVALRSPIERKNVFN 240
QY 241 TYVAVAASVVTALISGSSLAHPQCKISKTYGHSAYLPEGVAVDTSCHLIPSPMLPIVLGLV 300
   |||||||
Db 241 TYVAVAASVVTALISGSSLAHPQCKISKTYGHSAYLPEGVAVDTSCHLIPSPMLPIVLGLV 300
QY 301 AGLISVYGAKYLPCCCNRVIGIPHSSTMGYNFSLGLLGEIITYVLVLDTVGAGNGMIG 360
   |||||||
Db 301 AGLISVYGAKYLPCCCNRVIGIPHSSTMGYNFSLGLLGEIITYVLVLDTVGAGNGMIG 360
QY 361 FOVLSTIGELSLAIVIALTSGLLTGLLNLIKIRKAPHEAKYFPDQVFWKPHLAVGF 417
   |||||||
Db 361 FOVLSTIGELSLAIVIALTSGLLTGLLNLIKIRKAPHEAKYFPDQVFWKPHLAVGF 417

RESULT 2
ID 090Q21 PRELIMINARY: PRT: 417 AA.
AC 090Q21:
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-JUN-2000 (Tremblrel. 14, Last annotation update)
DE RH BLOOD GROUP D ANTIGEN (RHD).
GN RHLYO.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TO:
RA Hyodo H., Ishikawa Y., Kashiwase K., Ogawa A., Watanabe Y.,
RA Tsuneyama H., Toyoda C., Uchikawa M., Akaza T., Fujii T.;
RT "Polymorphisms of RhDya in Japanese."
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB018968; BAA81901.1;
DR InterPro: IPR001905;
DR InterPro: IPR002229;
DR Pfam: PF00909; Ammonium_transp; 1.
DR PRINTS: PR00342; RHESUSRHD.
SQ SEQUENCE 417 AA; 45177 MW; B72327B16747E6ID CRC64;

Query Match 93.0%; Score 1976; DB 4; Length 417;
Best Local Similarity 94.7%; Pred. No. 6.3e-118;
Matches 395; Conservative 6; Mismatches 16; Indels 0; Gaps 0;

QY 1 MSCGPRSVGRCLPLCALTEALILILEFFTHYDASLEDDKGLVASVOGODLTWMAI 60
   |||||||
Db 1 MSCGPRSVGRCLPLCALTEALILILEFFTHYDASLEDDKGLVASVOGODLTWMAI 60
QY 61 GLGFLTSSFRHSSWSVAFLMFLALGVQWAILDGLFSOPPSGKVYITLFSIWLATMSA 120
   |||||||
Db 61 GLGFLTSSFRHSSWSVAFLMFLALGVQWAILDGLFSOPPSGKVYITLFSIWLATMSA 120
QY 121 LSVLISYDAVLGKYNLAQLVVMVLEVTALGNLRMVISNIFNTDYHMMMHIIYFAAYFG 180
   |||||||
Db 121 LSVLISYDAVLGKYNLAQLVVMVLEVTALGNLRMVISNIFNTDYHMMMHIIYFAAYFG 180
QY 181 LTVAMCPLPKPLPECTEDDDQATIPSLSAMGLFLMFPVSVALRSPIERKNVFN 240
   |||||||
Db 181 LTVAMCPLPKPLPECTEDDDQATIPSLSAMGLFLMFPVSVALRSPIERKNVFN 240
QY 241 TYVAVAASVVTALISGSSLAHPQCKISKTYGHSAYLPEGVAVDTSCHLIPSPMLPIVLGLV 300
   |||||||
Db 241 TYVAVAASVVTALISGSSLAHPQCKISKTYGHSAYLPEGVAVDTSCHLIPSPMLPIVLGLV 300
QY 301 AGLISVYGAKYLPCCCNRVIGIPHSSTMGYNFSLGLLGEIITYVLVLDTVGAGNGMIG 360
   |||||||
Db 301 AGLISVYGAKYLPCCCNRVIGIPHSSTMGYNFSLGLLGEIITYVLVLDTVGAGNGMIG 360

```

```

Db 301 AGLISVYGAKYLPCCCNRVIGIPHSSTMGYNFSLGLLGEIITYVLVLDTVGAGNGMIG 360
QY 361 FOVLSTIGELSLAIVIALTSGLLTGLLNLIKIRKAPHEAKYFPDQVFWKPHLAVGF 417
   |||||||
Db 361 FOVLSTIGELSLAIVIALTSGLLTGLLNLIKIRKAPHEAKYFPDQVFWKPHLAVGF 417

RESULT 3
ID 090Q20 PRELIMINARY: PRT: 417 AA.
AC 090Q20:
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-JUN-2000 (Tremblrel. 14, Last annotation update)
DE RH BLOOD GROUP D ANTIGEN (RHD).
GN RHLYO.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TO:
RA Hyodo H., Ishikawa Y., Kashiwase K., Ogawa A., Watanabe Y.,
RA Tsuneyama H., Toyoda C., Uchikawa M., Akaza T., Fujii T.;
RT "Polymorphisms of RhDya in Japanese."
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB018969; BAA82159.1;
DR InterPro: IPR001905;
DR InterPro: IPR002229;
DR Pfam: PF00909; Ammonium_transp; 1.
DR PRINTS: PR00342; RHESUSRHD.
SQ SEQUENCE 417 AA; 45179 MW; 62B1DD24064F7D5 CRC64;

Query Match 93.0%; Score 1976; DB 4; Length 417;
Best Local Similarity 95.0%; Pred. No. 6.3e-118;
Matches 396; Conservative 4; Mismatches 17; Indels 0; Gaps 0;

QY 1 MSCGPRSVGRCLPLCALTEALILILEFFTHYDASLEDDKGLVASVOGODLTWMAI 60
   |||||||
Db 1 MSCGPRSVGRCLPLCALTEALILILEFFTHYDASLEDDKGLVASVOGODLTWMAI 60
QY 61 GLGFLTSSFRHSSWSVAFLMFLALGVQWAILDGLFSOPPSGKVYITLFSIWLATMSA 120
   |||||||
Db 61 GLGFLTSSFRHSSWSVAFLMFLALGVQWAILDGLFSOPPSGKVYITLFSIWLATMSA 120
QY 121 LSVLISYDAVLGKYNLAQLVVMVLEVTALGNLRMVISNIFNTDYHMMMHIIYFAAYFG 180
   |||||||
Db 121 LSVLISYDAVLGKYNLAQLVVMVLEVTALGNLRMVISNIFNTDYHMMMHIIYFAAYFG 180
QY 181 LTVAMCPLPKPLPECTEDDDQATIPSLSAMGLFLMFPVSVALRSPIERKNVFN 240
   |||||||
Db 181 LTVAMCPLPKPLPECTEDDDQATIPSLSAMGLFLMFPVSVALRSPIERKNVFN 240
QY 241 TYVAVAASVVTALISGSSLAHPQCKISKTYGHSAYLPEGVAVDTSCHLIPSPMLPIVLGLV 300
   |||||||
Db 241 TYVAVAASVVTALISGSSLAHPQCKISKTYGHSAYLPEGVAVDTSCHLIPSPMLPIVLGLV 300
QY 301 AGLISVYGAKYLPCCCNRVIGIPHSSTMGYNFSLGLLGEIITYVLVLDTVGAGNGMIG 360
   |||||||
Db 301 AGLISVYGAKYLPCCCNRVIGIPHSSTMGYNFSLGLLGEIITYVLVLDTVGAGNGMIG 360
QY 361 FOVLSTIGELSLAIVIALTSGLLTGLLNLIKIRKAPHEAKYFPDQVFWKPHLAVGF 417
   |||||||
Db 361 FOVLSTIGELSLAIVIALTSGLLTGLLNLIKIRKAPHEAKYFPDQVFWKPHLAVGF 417

```

```

RESULT 4
ID 090K19 PRELIMINARY: PRT: 417 AA.
AC 090K19:
DT 01-MAY-2000 (Tremblrel. 13, Created)

Query Match 93.0%; Score 1976; DB 4; Length 417;
Best Local Similarity 95.0%; Pred. No. 6.3e-118;
Matches 396; Conservative 4; Mismatches 17; Indels 0; Gaps 0;

QY 1 MSCGPRSVGRCLPLCALTEALILILEFFTHYDASLEDDKGLVASVOGODLTWMAI 60
   |||||||
Db 1 MSCGPRSVGRCLPLCALTEALILILEFFTHYDASLEDDKGLVASVOGODLTWMAI 60
QY 61 GLGFLTSSFRHSSWSVAFLMFLALGVQWAILDGLFSOPPSGKVYITLFSIWLATMSA 120
   |||||||
Db 61 GLGFLTSSFRHSSWSVAFLMFLALGVQWAILDGLFSOPPSGKVYITLFSIWLATMSA 120
QY 121 LSVLISYDAVLGKYNLAQLVVMVLEVTALGNLRMVISNIFNTDYHMMMHIIYFAAYFG 180
   |||||||
Db 121 LSVLISYDAVLGKYNLAQLVVMVLEVTALGNLRMVISNIFNTDYHMMMHIIYFAAYFG 180
QY 181 LTVAMCPLPKPLPECTEDDDQATIPSLSAMGLFLMFPVSVALRSPIERKNVFN 240
   |||||||
Db 181 LTVAMCPLPKPLPECTEDDDQATIPSLSAMGLFLMFPVSVALRSPIERKNVFN 240
QY 241 TYVAVAASVVTALISGSSLAHPQCKISKTYGHSAYLPEGVAVDTSCHLIPSPMLPIVLGLV 300
   |||||||
Db 241 TYVAVAASVVTALISGSSLAHPQCKISKTYGHSAYLPEGVAVDTSCHLIPSPMLPIVLGLV 300
QY 301 AGLISVYGAKYLPCCCNRVIGIPHSSTMGYNFSLGLLGEIITYVLVLDTVGAGNGMIG 360
   |||||||
Db 301 AGLISVYGAKYLPCCCNRVIGIPHSSTMGYNFSLGLLGEIITYVLVLDTVGAGNGMIG 360
QY 361 FOVLSTIGELSLAIVIALTSGLLTGLLNLIKIRKAPHEAKYFPDQVFWKPHLAVGF 417
   |||||||
Db 361 FOVLSTIGELSLAIVIALTSGLLTGLLNLIKIRKAPHEAKYFPDQVFWKPHLAVGF 417

```

```

DT 01-MAY-2000 (Tremblrel. 13, last sequence update)
DE 01-JUN-2000 (Tremblrel. 14, last annotation update)
DE RHD TYPE IITA PROTEIN.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97398395; PubMed=9256293;
RA Huang C.H., Chen Y., Reid M.;
RT "human D(IIa) erythrocytes: Rhd protein is associated with multiple
RT dispersed amino acid variations.";
RL Am. J. Hematol. 55:139-145(1997).
RN [2]
RP SEQUENCE FROM N.A.
RA Huang C.H.;
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF187846; AAF04565.1; -
DR InterPro; IPR001905; -
DR InterPro; IPR002229; -
DR Pfam; PF00909; Ammonium transp. 1.
DR PRINTS; PR00342; RHESUSRHD.
SQ SEQUENCE 417 AA; 45273 MW; 491FBIIE1BA37957 CRC64;

Query Match          92.9%; Score 1975; DB 4; Length 417;
Best Local Similarity 95.0%; Pred. No. 7.3e-118;
Matches 396; Conservative 4; Mismatches 17; Indels 0; Gaps 0;

QY 1 MSCKYPSVGRCLPLCALTEALILLFYFTHYDASLEQKGLVASYQVGDLTVAAI 60
DB 1 MSKYPSVGRCLPLMALTEALILLFYFTHYDASLEQKGLVASYQVGDLTVAAL 60

QY 61 GLGLTSSFRHSSVAFLFMLALGVOMAILLDGFLSQPPSGKVITLFSRLATMSA 120
DB 61 GLGLTSSFRHSSVAFLFMLALGVOMAILLDGFLSQPPSGKVITLFSRLATMSA 120

QY 121 LSVLISVDAYLGKYNLAQLVVMVLEVTGLGNLBMVISNIFNFDYHNMHHIYFAAYFG 180
DB 121 LSVLISVDAYLGKYNLAQLVVMVLEVTGLGNLBMVISNIFNFDYHNMHHIYFAAYFG 180

QY 121 LSVLISVDAYLGKYNLAQLVVMVLEVTGLGNLBMVISNIFNFDYHNMHHIYFAAYFG 180
DB 121 LSVLISVDAYLGKYNLAQLVVMVLEVTGLGNLBMVISNIFNFDYHNMHHIYFAAYFG 180

QY 181 LTVAMCLPKPLPECTEDNDQFATIPSLAMGALFLMFPSPVSALLRPIERKNVFN 240
DB 181 LTVAMCLPKPLPECTEDNDQFATIPSLAMGALFLMFPSPVSALLRPIERKNVFN 240

QY 241 TYAVAVSVTAISGSSLAHPQGIKSTYGHSAVLPEGVAVDTSCHLIPSPWLPVIGLV 300
DB 241 TYAVAVSVTAISGSSLAHPQGIKSTYGHSAVLPEGVAVDTSCHLIPSPWLPVIGLV 300

QY 301 AGLISV-GAKYLPCCCNRVLGIPHSITMGVNFSLGLLEITIVYLLVDLVGANGMIG 360
DB 301 AGLISVGGAKYLPCCCNRVLGIPHSITMGVNFSLGLLEITIVYLLVDLVGANGMIG 360

QY 361 FOVLISIGELSLAIVIALTSGLTALLNLKIRAPHAKEFFDQVWPKPHLAVGF 417
DB 361 FOVLISIGELSLAIVIALTSGLTALLNLKIRAPHAKEFFDQVWPKPHLAVGF 417

RESULT 5
Q9U022 PRELIMINARY; PRT; 417 AA.
AC Q9U022;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, last sequence update)
DT 01-JUN-2000 (Tremblrel. 14, last annotation update)
DE RH BLOOD GROUP D ANTIGEN (RHD).
GN RHDVA.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]

```

```

RP SEQUENCE FROM N.A.
RC STRAIN=TO;
RA Hyodo H., Ishikawa Y., Kashiwase K., Ogawa A., Watanabe Y.,
RA Tsuneyama H., Toyoda C., Uchikawa M., Akaza T., Fujii T.;
RA "Polymorphisms of Rhdv in Japanese.";
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB018967; BA681900.1; -
DR InterPro; IPR001905; -
DR InterPro; IPR002229; -
DR Pfam; PF00909; Ammonium transp. 1.
DR PRINTS; PR00342; RHESUSRHD.
SQ SEQUENCE 417 AA; 45211 MW; 764B092CBF95ECC4 CRC64;

Query Match          92.9%; Score 1974; DB 4; Length 417;
Best Local Similarity 94.7%; Pred. No. 8.5e-118;
Matches 395; Conservative 5; Mismatches 17; Indels 0; Gaps 0;

QY 1 MSCKYPSVGRCLPLCALTEALILLFYFTHYDASLEQKGLVASYQVGDLTVAAI 60
DB 1 MSKYPSVGRCLPLMALTEALILLFYFTHYDASLEQKGLVASYQVGDLTVAAI 60

QY 61 GLGLTSSFRHSSVAFLFMLALGVOMAILLDGFLSQPPSGKVITLFSRLATMSA 120
DB 61 GLGLTSSFRHSSVAFLFMLALGVOMAILLDGFLSQPPSGKVITLFSRLATMSA 120

QY 121 LSVLISVDAYLGKYNLAQLVVMVLEVTGLGNLBMVISNIFNFDYHNMHHIYFAAYFG 180
DB 121 LSVLISVDAYLGKYNLAQLVVMVLEVTGLGNLBMVISNIFNFDYHNMHHIYFAAYFG 180

QY 181 LTVAMCLPKPLPECTEDNDQFATIPSLAMGALFLMFPSPVSALLRPIERKNVFN 240
DB 181 LTVAMCLPKPLPECTEDNDQFATIPSLAMGALFLMFPSPVSALLRPIERKNVFN 240

QY 241 TYAVAVSVTAISGSSLAHPQGIKSTYGHSAVLPEGVAVDTSCHLIPSPWLPVIGLV 300
DB 241 TYAVAVSVTAISGSSLAHPQGIKSTYGHSAVLPEGVAVDTSCHLIPSPWLPVIGLV 300

QY 301 AGLISV-GAKYLPCCCNRVLGIPHSITMGVNFSLGLLEITIVYLLVDLVGANGMIG 360
DB 301 AGLISVGGAKYLPCCCNRVLGIPHSITMGVNFSLGLLEITIVYLLVDLVGANGMIG 360

QY 361 FOVLISIGELSLAIVIALTSGLTALLNLKIRAPHAKEFFDQVWPKPHLAVGF 417
DB 361 FOVLISIGELSLAIVIALTSGLTALLNLKIRAPHAKEFFDQVWPKPHLAVGF 417

RESULT 6
Q9NXY8 PRELIMINARY; PRT; 417 AA.
AC Q9NXY8;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, last sequence update)
DT 01-MAR-2001 (Tremblrel. 16, last annotation update)
DE RHESUS BLOOD GROUP, D ANTIGEN.
GN RHD.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Wagner F.F., Flegel W.A.;
RT "Molecular structure of weak D type 18, another RHD allele detected
RT among weak D phenotypes.";
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ287289; CAB75731.1; -
DR EMBL; AJ287290; CAB75731.1; JOINED.
DR EMBL; AJ287291; CAB75731.1; JOINED.
DR EMBL; AJ287292; CAB75731.1; JOINED.
DR EMBL; AJ287293; CAB75731.1; JOINED.
DR EMBL; AJ287294; CAB75731.1; JOINED.
DR EMBL; AJ287295; CAB75731.1; JOINED.

```



```

Db 181 LSVAMCLPKPLPECTEDKDDQTATIPSLSAMGLFLMFWPSPSALLRSPRIERKNAFVN 240
QY 241 TYVAVAASVTAIAGSSSLAHPOGKISKTYGSAVLPEGVAVGTSCHLIPSPMLPIVGLV 300
Db 241 TYVAVAASVTAIAGSSSLAHPOGKISKTYGSAVLPEGVAVGTSCHLIPSPMLPIVGLV 300
QY 301 AGLISVGAAYLPCCCNRVIGIPHSISMGYNFSLGLLGLIETIYVLLVDTVGANGMIG 360
Db 301 AGLISVGAAYLPCCCNRVIGIPHSISMGYNFSLGLLGLIETIYVLLVDTVGANGMIG 360
QY 361 FOVLISIGELSLAIVIALTSGLLTALLNLKIRAPHEAKYFDDQVFWKPEPFLAVGF 417
Db 361 FOVLISIGELSLAIVIALTSGLLTALLNLKIRAPHEAKYFDDQVFWKPEPFLAVGF 417

RESULT 9
ID Q9H248 PRELIMINARY; PRT; 417 AA.
AC Q9H248;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DE 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE RHESUS D CATEGORY VI TYPE IV PROTEIN.
GN RHD.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Esteban R., Nogues N., Montero R., Hurtado M., Maroto S., Ribera A.;
RT "Characterization of a new RHD category VI (type IV) allele,
RT predominant in the Spanish population.";
RL Submitted (OCT-2000) to the EMBL/GenBank/DBD databases.
DR EMBL; AF312679; AAC40872.1; -;
SQ SEQUENCE 417 AA; 45293 MW; 03FD058960E41D89 CRC64;

Query Match 91.4%; Score 1942; DB 4; Length 417;
Best Local Similarity 93.0%; Pred. No. 9e-116;
Matches 388; Conservative 8; Mismatches 21; Indels 0; Gaps 0;

QY 1 MSCKYPSVGRCLPLCALTEALILLFYFTHYDASLEDDKGLVASVYOGDULTVMAAI 60
Db 1 MSCKYPSVGRCLPLCALTEALILLFYFTHYDASLEDDKGLVASVYOGDULTVMAAI 60
QY 61 GLGFLTSSFRHRHSSVAFNLFMLALGVOMAILLDGFLSQPSPGKVVITLFSILATMSA 120
Db 61 GLGFLTSSFRHRHSSVAFNLFMLALGVOMAILLDGFLSQPSPGKVVITLFSILATMSA 120
QY 121 LSVLISDAVLGKYNLAOLVYVVEVTGLGNLHMVINSIFNTDYHNMHMIYFAAYFG 180
Db 121 LSVLISDAVLGKYNLAOLVYVVEVTGLGNLHMVINSIFNTDYHNMHMIYFAAYFG 180
QY 121 MSVLISGAVLGKYNLAOLVYVVEVTGLGNLHMVINSIFNTDYHNMHMIYFAAYFG 180
Db 121 MSVLISGAVLGKYNLAOLVYVVEVTGLGNLHMVINSIFNTDYHNMHMIYFAAYFG 180
QY 181 LTVAMCLPKPLPECTEDDQATIPSLSAMGLFLMFWPSPSALLRSPRIERKNAFVN 240
Db 181 LTVAMCLPKPLPECTEDDQATIPSLSAMGLFLMFWPSPSALLRSPRIERKNAFVN 240
QY 241 TYVAVAASVTAIAGSSSLAHPOGKISKTYGSAVLPEGVAVGTSCHLIPSPMLPIVGLV 300
Db 241 TYVAVAASVTAIAGSSSLAHPOGKISKTYGSAVLPEGVAVGTSCHLIPSPMLPIVGLV 300
QY 301 AGLISVGAAYLPCCCNRVIGIPHSISMGYNFSLGLLGLIETIYVLLVDTVGANGMIG 360
Db 301 AGLISVGAAYLPCCCNRVIGIPHSISMGYNFSLGLLGLIETIYVLLVDTVGANGMIG 360
QY 361 FOVLISIGELSLAIVIALTSGLLTALLNLKIRAPHEAKYFDDQVFWKPEPFLAVGF 417
Db 361 FOVLISIGELSLAIVIALTSGLLTALLNLKIRAPHEAKYFDDQVFWKPEPFLAVGF 417

RESULT 10
Q9UD25

```

```

ID Q9UD25 PRELIMINARY; PRT; 417 AA.
AC Q9UD25;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE RHESUS D CATEGORY VI TYPE III PROTEIN.
GN RHD.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BLOOD;
RX MEDLINE=98158656; Pubmed=9490704;
RA Wagner F.F., Gassner C., Mueller T.H., Schoenitzer D., Schunter F.,
RA Flegel W.A.;
RT "Three molecular structures cause Rhesus D category VI phenotypes with
RT distinct immunohematologic features.";
RL Blood 91:2157-2168(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=BLOOD;
RX MEDLINE=94245182; Pubmed=8188244;
RA Cherif-Zahar B., Le Van Kim C., Rouillac C., Raynal V., Carton J.,
RA Colin Y.;
RT "Organization of the gene (RHCE) encoding the human blood group Rhce
RT antigens and characterization of the promoter region.";
RL Genomics 19:68-74(1994).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=BLOOD;
RX MEDLINE=9306356; Pubmed=1438298;
RA Le Van Kim C., Mourio I., Cherif-Zahar B., Raynal V., Cherrier C.,
RA Carton J.P., Colin Y.;
RT "Molecular cloning and primary structure of the human blood group RHD
RT polypeptide.";
RL Proc. Natl. Acad. Sci. U.S.A. 89:10925-10929(1992).
DR EMBL; 297026; CAB09722.1; -;
DR InterPro; IPR001905; -;
DR InterPro; IPR002229; -;
DR Pfam; PF00909; Ammonium_transp; 1.
DR PRINTS; PR00342; RHESUSRHD.
SQ SEQUENCE 417 AA; 45247 MW; 1887A3C519149E8F CRC64;

Query Match 90.9%; Score 1932; DB 4; Length 417;
Best Local Similarity 92.6%; Pred. No. 3.9e-115;
Matches 386; Conservative 9; Mismatches 22; Indels 0; Gaps 0;

QY 1 MSCKYPSVGRCLPLCALTEALILLFYFTHYDASLEDDKGLVASVYOGDULTVMAAI 60
Db 1 MSCKYPSVGRCLPLCALTEALILLFYFTHYDASLEDDKGLVASVYOGDULTVMAAI 60
QY 61 GLGFLTSSFRHRHSSVAFNLFMLALGVOMAILLDGFLSQPSPGKVVITLFSILATMSA 120
Db 61 GLGFLTSSFRHRHSSVAFNLFMLALGVOMAILLDGFLSQPSPGKVVITLFSILATMSA 120
QY 121 LSVLISDAVLGKYNLAOLVYVVEVTGLGNLHMVINSIFNTDYHNMHMIYFAAYFG 180
Db 121 LSVLISDAVLGKYNLAOLVYVVEVTGLGNLHMVINSIFNTDYHNMHMIYFAAYFG 180
QY 121 MSVLISGAVLGKYNLAOLVYVVEVTGLGNLHMVINSIFNTDYHNMHMIYFAAYFG 180
Db 121 MSVLISGAVLGKYNLAOLVYVVEVTGLGNLHMVINSIFNTDYHNMHMIYFAAYFG 180
QY 181 LTVAMCLPKPLPECTEDDQATIPSLSAMGLFLMFWPSPSALLRSPRIERKNAFVN 240
Db 181 LTVAMCLPKPLPECTEDDQATIPSLSAMGLFLMFWPSPSALLRSPRIERKNAFVN 240
QY 241 TYVAVAASVTAIAGSSSLAHPOGKISKTYGSAVLPEGVAVGTSCHLIPSPMLPIVGLV 300
Db 241 TYVAVAASVTAIAGSSSLAHPOGKISKTYGSAVLPEGVAVGTSCHLIPSPMLPIVGLV 300
QY 301 AGLISVGAAYLPCCCNRVIGIPHSISMGYNFSLGLLGLIETIYVLLVDTVGANGMIG 360
Db 301 AGLISVGAAYLPCCCNRVIGIPHSISMGYNFSLGLLGLIETIYVLLVDTVGANGMIG 360

```

OY 361 FOYLLSIGELSLAIVALTSGLLTALLNLKIKAPHEAKYFDDQVWKPRLAVGF 417
 Db 361 FOYLLSIGELSLAIVALTSGLLTALLNLKIKAPHEAKYFDDQVWKPRLAVGF 417

RESULT 11
 ID 09UK74 PRELIMINARY; PRT; 417 AA.

AC 09UK74;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
 DE RH BLOOD GROUP ANTIGEN EVANS.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;

RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-PERIPHERAL BLOOD;
 RX MEDLINE=96404415; Pubmed=8808597;
 RA Huang C.H., Chen Y., Reid M., Ghosh S.;
 RT "Genetic recombination at the human RH locus: a family study of the
 RT red-cell Evans phenotype reveals a transfer of exons 2-6 from the RHD
 RT to the RHCE gene."
 RL Am. J. Hum. Genet. 59:825-833(1996).
 RN [2]

RP SEQUENCE FROM N.A.
 RC TISSUE-PERIPHERAL BLOOD;
 RA Huang C.H.;
 RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF17938; AAD5773.1; -
 DR InterPro: IPR001762; -
 DR InterPro: IPR001905; -
 DR InterPro: IPR002229; -
 DR Pfam: PF00909; Ammonium_transp; 1.
 DR PRINTS: PR00342; RHESUSRD.
 DR Prodom: PD000664; -; 1.
 SO SEQUENCE 417 AA; 45445 MW; D8AADB148ADC73F7 CRC64;

Query Match 89.3%; Score 1897; DB 4; Length 417;
 Best Local Similarity 92.1%; Pred. No. 6.4e-113;
 Matches 384; Conservative 5; Mismatches 28; Indels 0; Gaps 0;

OY 1 MSCKYPRSVGRCLPLCALTEALILFFFTHYDASLEQKGLVASYOGODLTWAAI 60
 Db 1 MSCKYPRSVGRCLPLCALTEALILFFFTHYDASLEQKGLVASYOGODLTWAAI 60
 OY 61 GLGFLTSSFRHSMSSVAFLNFMALGVOMAILDGLSQPSGKVITLFSIATMSA 120
 Db 61 GLGFLTSSFRHSMSSVAFLNFMALGVOMAILDGLSQPSGKVITLFSIATMSA 120
 OY 121 LSVLISVDALGVNLAQLVVMVLEVTdLGNLRMVISNIFNTDYHNMNHIYFAAYFG 180
 Db 121 LSVLISVDALGVNLAQLVVMVLEVTdLGNLRMVISNIFNTDYHNMNHIYFAAYFG 180
 OY 181 LTVAMCLPKPLPEGTEDNDQRTATIPSLAAMLGALFLMFPVSNSALLSPERKNAVEN 240
 Db 181 LTVAMCLPKPLPEGTEDNDQRTATIPSLAAMLGALFLMFPVSNSALLSPERKNAVEN 240
 OY 181 LTVAMCLPKPLPEGTEDNDQRTATIPSLAAMLGALFLMFPVSNSALLSPERKNAVEN 240
 Db 181 LTVAMCLPKPLPEGTEDNDQRTATIPSLAAMLGALFLMFPVSNSALLSPERKNAVEN 240
 OY 241 TYAANAVSVTAISGSSLAHPQCKISKTYGHSANVLPegVAVGTSCHLIPSPMLpIVLGLV 300
 Db 241 TYAANAVSVTAISGSSLAHPQCKISKTYGHSANVLPegVAVGTSCHLIPSPMLpIVLGLV 300
 OY 301 AGLISYGAAYLPGCCNRVLGIPHSIMGYNSFLGLLEIYIVLLVDTVGAGNGMIG 360
 Db 301 AGLISYGAAYLPGCCNRVLGIPHSIMGYNSFLGLLEIYIVLLVDTVGAGNGMIG 360
 OY 361 FOYLLSIGELSLAIVALTSGLLTALLNLKIKAPHEAKYFDDQVWKPRLAVGF 417
 Db 361 FOYLLSIGELSLAIVALTSGLLTALLNLKIKAPHEAKYFDDQVWKPRLAVGF 417

RESULT 12
 ID 09NPKO PRELIMINARY; PRT; 417 AA.

AC 09NPKO;
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
 DE RH BLOOD GROUP ANTIGEN RHD.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;

RN [1]
 RP SEQUENCE FROM N.A.
 RA Hyodo H., Ishikawa Y., Tsuneyama H., Kashiwase K., Toyoda C.,
 RA Uchikawa M., Akaza T., Fujii T., Kozuma S., Taketani Y., Juji T.;
 RT "New Rhivb identified in Japanese."
 RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AB037270; BAA90298.1; -
 DR InterPro: IPR001905; -
 DR InterPro: IPR002229; -
 DR Pfam: PF00909; Ammonium_transp; 1.
 DR PRINTS: PR00342; RHESUSRD.
 SO SEQUENCE 417 AA; 45429 MW; 647D4A7678CB498 CRC64;

Query Match 89.2%; Score 1896; DB 4; Length 417;
 Best Local Similarity 92.3%; Pred. No. 7.4e-113;
 Matches 385; Conservative 4; Mismatches 28; Indels 0; Gaps 0;

OY 1 MSCKYPRSVGRCLPLCALTEALILFFFTHYDASLEQKGLVASYOGODLTWAAI 60
 Db 1 MSCKYPRSVGRCLPLCALTEALILFFFTHYDASLEQKGLVASYOGODLTWAAI 60
 OY 61 GLGFLTSSFRHSMSSVAFLNFMALGVOMAILDGLSQPSGKVITLFSIATMSA 120
 Db 61 GLGFLTSSFRHSMSSVAFLNFMALGVOMAILDGLSQPSGKVITLFSIATMSA 120
 OY 121 LSVLISVDALGVNLAQLVVMVLEVTdLGNLRMVISNIFNTDYHNMNHIYFAAYFG 180
 Db 121 LSVLISVDALGVNLAQLVVMVLEVTdLGNLRMVISNIFNTDYHNMNHIYFAAYFG 180
 OY 181 LTVAMCLPKPLPEGTEDNDQRTATIPSLAAMLGALFLMFPVSNSALLSPERKNAVEN 240
 Db 181 LTVAMCLPKPLPEGTEDNDQRTATIPSLAAMLGALFLMFPVSNSALLSPERKNAVEN 240
 OY 181 LTVAMCLPKPLPEGTEDNDQRTATIPSLAAMLGALFLMFPVSNSALLSPERKNAVEN 240
 Db 181 LTVAMCLPKPLPEGTEDNDQRTATIPSLAAMLGALFLMFPVSNSALLSPERKNAVEN 240
 OY 241 TYAANAVSVTAISGSSLAHPQCKISKTYGHSANVLPegVAVGTSCHLIPSPMLpIVLGLV 300
 Db 241 TYAANAVSVTAISGSSLAHPQCKISKTYGHSANVLPegVAVGTSCHLIPSPMLpIVLGLV 300
 OY 301 AGLISYGAAYLPGCCNRVLGIPHSIMGYNSFLGLLEIYIVLLVDTVGAGNGMIG 360
 Db 301 AGLISYGAAYLPGCCNRVLGIPHSIMGYNSFLGLLEIYIVLLVDTVGAGNGMIG 360
 OY 361 FOYLLSIGELSLAIVALTSGLLTALLNLKIKAPHEAKYFDDQVWKPRLAVGF 417
 Db 361 FOYLLSIGELSLAIVALTSGLLTALLNLKIKAPHEAKYFDDQVWKPRLAVGF 417

RESULT 13
 ID 09UPNO PRELIMINARY; PRT; 417 AA.
 AC 09UPNO;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
 DE RH BLOOD CE GROUP ANTIGEN POLYPEPTIDE.
 GN RHCE.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 12, 2001, 07:33:15 ; Search time 20.48 Seconds
(without alignments)
1237.345 Million cell updates/sec

Title: 09-600714-2
Perfect score: 2125
Sequence: 1 MSCKRPSVGRCLPLCALTL.....AKYFDQYFWKPHLAVGFX 418

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 412676 seqs, 60623988 residues
Total number of hits satisfying chosen parameters: 412676

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-Processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

A.Geneseq_0601:*

1:	/SIDSB/gcgdata/geneseq/geneseqp/AA1980.DAT:*
2:	/SIDSB/gcgdata/geneseq/geneseqp/AA1981.DAT:*
3:	/SIDSB/gcgdata/geneseq/geneseqp/AA1982.DAT:*
4:	/SIDSB/gcgdata/geneseq/geneseqp/AA1983.DAT:*
5:	/SIDSB/gcgdata/geneseq/geneseqp/AA1984.DAT:*
6:	/SIDSB/gcgdata/geneseq/geneseqp/AA1985.DAT:*
7:	/SIDSB/gcgdata/geneseq/geneseqp/AA1986.DAT:*
8:	/SIDSB/gcgdata/geneseq/geneseqp/AA1987.DAT:*
9:	/SIDSB/gcgdata/geneseq/geneseqp/AA1988.DAT:*
10:	/SIDSB/gcgdata/geneseq/geneseqp/AA1989.DAT:*
11:	/SIDSB/gcgdata/geneseq/geneseqp/AA1990.DAT:*
12:	/SIDSB/gcgdata/geneseq/geneseqp/AA1991.DAT:*
13:	/SIDSB/gcgdata/geneseq/geneseqp/AA1992.DAT:*
14:	/SIDSB/gcgdata/geneseq/geneseqp/AA1993.DAT:*
15:	/SIDSB/gcgdata/geneseq/geneseqp/AA1994.DAT:*
16:	/SIDSB/gcgdata/geneseq/geneseqp/AA1995.DAT:*
17:	/SIDSB/gcgdata/geneseq/geneseqp/AA1996.DAT:*
18:	/SIDSB/gcgdata/geneseq/geneseqp/AA1997.DAT:*
19:	/SIDSB/gcgdata/geneseq/geneseqp/AA1998.DAT:*
20:	/SIDSB/gcgdata/geneseq/geneseqp/AA1999.DAT:*
21:	/SIDSB/gcgdata/geneseq/geneseqp/AA2000.DAT:*
22:	/SIDSB/gcgdata/geneseq/geneseqp/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1982	93.3	417	21	AA03341
2	1980	93.2	417	20	AA24056
3	1859	87.5	417	21	AA03339
4	1854	87.2	417	21	AA03338
5	1838	86.5	417	21	AA03342
6	1833	86.3	417	21	AA03340
7	497	23.4	473	21	AA042135
8	497	23.4	479	21	AA029656
9	142	6.7	308	21	AA086502
10	109.5	5.2	495	21	AA039787
11	109.5	5.2	577	21	AA039786

12	104.5	4.9	449	19	AA098770	H. pylori GHPO 113
13	104.5	4.9	497	18	AA020821	H. pylori cell env
14	104	4.9	449	21	AA081556	Streptococcus pneu
15	104	4.9	465	22	AA078999	C. glutamicum SRT
16	102.5	4.8	449	19	AA011065	H. pylori ORF 09cp
17	102.5	4.8	1191	21	AA046051	Arabidopsis thalia
18	102.5	4.8	1195	21	AA046050	Arabidopsis thalia
19	102.5	4.8	1325	21	AA046049	Arabidopsis thalia
20	101.5	4.8	461	20	AA038791	Neisseria meningit
21	101	4.8	181	21	AA086279	Human secreted pro
22	100.5	4.7	461	20	AA038792	Neisseria meningit
23	100.5	4.7	758	22	AA079584	Cornebacterium gl
24	100	4.7	348	20	AA019899	B. burgdorferi ant
25	100	4.7	383	20	AA019898	B. burgdorferi ant
26	100	4.7	437	21	AA093270	Amino acid sequenc
27	100	4.7	565	22	AA076817	Cornebacterium gl
28	99	4.7	439	21	AA049911	Arabidopsis thalia
29	99	4.7	442	21	AA013571	Arabidopsis thalia
30	99	4.7	442	21	AA017244	Arabidopsis thalia
31	99	4.7	442	21	AA045471	Arabidopsis thalia
32	99	4.7	442	21	AA045499	Arabidopsis thalia
33	99	4.7	521	21	AA049910	Arabidopsis thalia
34	99	4.7	672	21	AA036170	Novel human transp
35	98.5	4.6	388	19	AA098618	H. pylori GHPO 231
36	98	4.6	405	22	AA079001	C. glutamicum SRT
37	98	4.6	437	21	AA093294	Amino acid sequenc
38	98	4.6	462	20	AA097820	Staphylococcus aur
39	97.5	4.6	437	21	AA049886	Arabidopsis thalia
40	97.5	4.6	476	21	AA070157	Staphylococcus aur
41	97.5	4.6	504	21	AA020282	Arabidopsis thalia
42	97.5	4.6	504	21	AA049885	Arabidopsis thalia
43	97.5	4.6	516	21	AA020281	Arabidopsis thalia
44	97.5	4.6	516	21	AA049884	Arabidopsis thalia
45	97.5	4.6	528	21	AA020280	Arabidopsis thalia

ALIGNMENTS

RESULT	1
AA03341	standard: Protein: 417 AA.
ID	AA03341
XX	
AC	AA03341:
XX	
DF	26-SEP-2000 (first entry)
XX	
DE	Human Rhd protein.
XX	
KW	Human: rhesus blood group system: Rh: Rhd; RHCE: sickle cell disease;
KM	thalassaemia: Rhc: Rho: Rhe: alloimmunisation prevention:
KW	autoimmune Rh haemolytic disease; rhesus protein; immunosuppressive;
XX	vaccine.
XX	
OS	Homo sapiens.
XX	
PN	W0200032632-A2.
XX	
PD	08-JUN-2000.
XX	
PF	01-DEC-1999; 99WO-GH04027.
XX	
PR	01-DEC-1998; 98GB-0026378.
XX	
PA	(UYAB-) UNIV ABERDEEN.
XX	(COMM-) COMMON SERVICES AGENCY SCOTTISH HEALTH S.
PI	Urbanlak SJ, Barker RN;
XX	
DR	WPI; 2000-412291/35.
XX	
PT	Composition for prevention of alloimmunization or immunosuppression of
XX	a response elicited by alloimmunization or an autoimmune hemolytic

PT disease, comprises an epitope of a rhesus protein

XX
PS Disclosure: Page 88-90; 92pp; English.

XX
CC Human blood contains the rhesus (Rh) blood group system, and humans can
CC either be RhD positive or negative. This can lead to complications
CC during transfusions or pregnancy if RhD negative individuals are exposed
CC to RhD positive blood, leading to them becoming immunised to produce
CC anti-D. The present invention relates to new human allo- and
CC auto-reactive T-cell epitopes (AA99760-Y99769 and AAB03201-B03337) from
CC RHD, RHC, RHE and RHF proteins. These epitopes bind to T-cells to
CC elicit an immune response, i.e. immunisation. These epitopes can be used
CC as a vaccine for the prevention of alloimmunisation or immunosuppression
CC of a response elicited by alloimmunisation or an autoimmune haemolytic
CC disease. Examples of autoimmune haemolytic diseases are sickle cell
CC disease and thalassemia. The present sequence is the human RhD protein
CC from which the epitopes in the present invention were derived.

XX
SQ Sequence 417 AA;

Query Match 93.3%; Score 1982; DB 21; Length 417;
Best Local Similarity 95.2%; Pred. No. 6.6e-185;
Matches 397; Conservative 4; Mismatches 16; Indels 0; Gaps 0;

QY 1 MSCKYPSVGRCLPLCALTEALILLFYFTHYDASLEDOKGLVASVYQGDLTWMAAI 60
1 msckypsvgrclplcaltealilfffthydasldeqkglvasyqgdltwmaai 60
DB 1 msckypsvgrclplcaltealilfffthydasldeqkglvasyqgdltwmaai 60

QY 61 GLGFLTSSFRHRSWSSVAENFLMALGVQMAILLDGLFSGPFGKVVITLFSTLWATMSA 120
61 glgfltsfrhrrswssvaenflmalgvqmailldglfsgpfgkvvitlfsrlatmsa 120
DB 61 glgfltsfrhrrswssvaenflmalgvqmailldglfsgpfgkvvitlfsrlatmsa 120

QY 121 LSVLISVDVAVLGKVNLAOLVVMVLEVTDLGNLRMVISNFFNPDYHNMNHITVFAAYFG 180
121 lsvlisdvavlgkvnlaolvvmvlevtdlgnlrmvisnffnptyhnmnhitvfaayfg 180
DB 121 lsvlisdvavlgkvnlaolvvmvlevtdlgnlrmvisnffnptyhnmnhitvfaayfg 180

QY 181 LTVAMCLPKPLPGTEEDNDQATIPSLSAMLGALFLMFEPSVNSALLRSPRIERKNAVFN 240
181 ltvamclpkplpgteedndqatipslsamlgalfmwfepsvnsallrsprierknafvn 240
DB 181 ltvamclpkplpgteedndqatipslsamlgalfmwfepsvnsallrsprierknafvn 240

QY 241 TYRAVAVSVYTAISGSSLAHPQCKISKTYGHSANVDPGVAVDTSCHLIPSPWIDVIGLV 300
241 tyraavsvytaisgsslahpqckisktyghsanvdpgevavdtschlipspwldvlglv 300
DB 241 tyraavsvytaisgsslahpqckisktyghsanvdpgevavdtschlipspwldvlglv 300

QY 301 AGLISVYGAKYLPGCCNRVIGIPHSSTMGYNFSLGLLEETIYVLLVLDTVGAGNGMIG 360
301 aglisygakylpgccnrvigiphsstmgynfslglleetiylvllvldtvgaqngmig 360
DB 301 aglisygakylpgccnrvigiphsstmgynfslglleetiylvllvldtvgaqngmig 360

QY 361 FOVLSTIGELSLAIVIALTSGLTALLNLKIRKAPHEAKYFDDQVFWKFPHLAVGF 417
361 fgvllstigelstlaivialtsglltllnllkirkapheakyyfddqvfwkfpahlavgf 417
DB 361 fgvllstigelstlaivialtsglltllnllkirkapheakyyfddqvfwkfpahlavgf 417

RESULT 2
AA924056 standard; Protein: 417 AA.
ID AA924056 standard; Protein: 417 AA.
XX
AC AA924056;
XX
DT 04-OCT-1999 (first entry)
XX
DE Protein encoded by the prevalent allele of the RhD gene.
XX
KM Allele; Rhesus D antigen; RhD; weak D phenotype; blood transfusion.
XX
OS Homo sapiens.
XX
PN WO9937763-A2.
XX
PD 29-JUL-1999.
XX
PF 18-DEC-1998; 98WO-EP08319.

XX
PR 23-JAN-1998; 98EP-0101203.
XX
PA (DRKB-) DRK BLUTSPENDEDIENST BADEN WUERTTEMBERG.
XX
PI Flegel WA, Wagner FF;
XX
DR WPI: 1999-469127/39.
XX
DR N-PSDB: AAX86522.
XX
PT Nucleic acid sequences correlated with Rhesus weak D phenotype,
PT useful for screening blood from donors and recipients for
PT transfusion methods
XX
PS Disclosure: Fig 2; 64pp; English.

XX
CC The present sequence is encoded by the prevalent allele of the Rhesus D
CC (RhD) antigen gene. The specification describes a RhD contributing to
CC or indicative of the weak D phenotype, where the RhD polynucleotide
CC carries at least one missense mutation as compared to the wild-type RhD,
CC in its transmembrane and/or intracellular regions, especially in amino
CC acid positions 2-16, 114-149, 179-225 or/and 267-397, with the proviso
CC that the D antigen does not carry a single missense mutation leading to
CC a F223V or T283I substitution. The probes and antibodies are useful in
CC the methods for detection of weak D phenotypes. Red blood cells, from
CC probands, are useful for the assessment of the affinity, avidity and/or
CC reactivity of monoclonal anti-D antibodies, polyclonal anti-D antisera or
CC of anti-globulin or anti-human-globulin antisera. Detecting the presence
CC of the RhD associated with weak D phenotype is useful for determining
CC that a patient in need of a blood transfusion is to be transfused with
CC RhD negative blood from a donor. Alternatively, testing for weak D
CC phenotype RhD in the blood of a donor is useful for determining whether
CC the donor blood should be excluded for transfusion to patients having
CC wild type RhD or weak D types, other than that of the donor weak D
CC type.

XX
SQ Sequence 417 AA;

Query Match 93.2%; Score 1980; DB 20; Length 417;
Best Local Similarity 95.2%; Pred. No. 1e-184;
Matches 397; Conservative 3; Mismatches 17; Indels 0; Gaps 0;

QY 1 MSCKYPSVGRCLPLCALTEALILLFYFTHYDASLEDOKGLVASVYQGDLTWMAAI 60
1 msckypsvgrclplcaltealilfffthydasldeqkglvasyqgdltwmaai 60
DB 1 msckypsvgrclplcaltealilfffthydasldeqkglvasyqgdltwmaai 60

QY 61 GLGFLTSSFRHRSWSSVAENFLMALGVQMAILLDGLFSGPFGKVVITLFSTLWATMSA 120
61 glgfltsfrhrrswssvaenflmalgvqmailldglfsgpfgkvvitlfsrlatmsa 120
DB 61 glgfltsfrhrrswssvaenflmalgvqmailldglfsgpfgkvvitlfsrlatmsa 120

QY 121 LSVLISVDVAVLGKVNLAOLVVMVLEVTDLGNLRMVISNFFNPDYHNMNHITVFAAYFG 180
121 lsvlisdvavlgkvnlaolvvmvlevtdlgnlrmvisnffnptyhnmnhitvfaayfg 180
DB 121 lsvlisdvavlgkvnlaolvvmvlevtdlgnlrmvisnffnptyhnmnhitvfaayfg 180

QY 181 LTVAMCLPKPLPGTEEDNDQATIPSLSAMLGALFLMFEPSVNSALLRSPRIERKNAVFN 240
181 ltvamclpkplpgteedndqatipslsamlgalfmwfepsvnsallrsprierknafvn 240
DB 181 ltvamclpkplpgteedndqatipslsamlgalfmwfepsvnsallrsprierknafvn 240

QY 241 TYRAVAVSVYTAISGSSLAHPQCKISKTYGHSANVDPGVAVDTSCHLIPSPWIDVIGLV 300
241 tyraavsvytaisgsslahpqckisktyghsanvdpgevavdtschlipspwldvlglv 300
DB 241 tyraavsvytaisgsslahpqckisktyghsanvdpgevavdtschlipspwldvlglv 300

QY 301 AGLISVYGAKYLPGCCNRVIGIPHSSTMGYNFSLGLLEETIYVLLVLDTVGAGNGMIG 360
301 aglisygakylpgccnrvigiphsstmgynfslglleetiylvllvldtvgaqngmig 360
DB 301 aglisygakylpgccnrvigiphsstmgynfslglleetiylvllvldtvgaqngmig 360

QY 361 FOVLSTIGELSLAIVIALTSGLTALLNLKIRKAPHEAKYFDDQVFWKFPHLAVGF 417
361 fgvllstigelstlaivialtsglltllnllkirkapheakyyfddqvfwkfpahlavgf 417
DB 361 fgvllstigelstlaivialtsglltllnllkirkapheakyyfddqvfwkfpahlavgf 417

RESULT 3

AAB03339
ID AAB03339 standard; Protein; 417 AA.
AC AAB03339;
XX
XX
DT 26-SEP-2000 (first entry)
DE Human RhCe protein.
XX
XX Human; rhesus blood group system; Rh; RhD; RhCE; sickle cell disease;
KM thalassaemia; Rhc; Rho; Rhe; Rhd; alloimmunisation prevention;
KW autoimmune Rh haemolytic disease; rhesus protein; immunosuppressive;
KW vaccine.
XX
XX Homo sapiens.
OS
XX
XX MO200032632-A2.
PN
XX
XX 08-JUN-2000.
PD
XX
XX 01-DEC-1999; 99MO-GB04027.
PE
XX
XX 01-DEC-1998; 98GB-0026378.
PR
XX
XX (UYAB-) UNIV ABERDEEN.
PA (COMM-) COMMON SERVICES AGENCY SCOTTISH HEALTH S.
XX
XX Urbanlak SJ, Barker RN;
PI
XX
XX WPI; 2000-412291/35.
DR
XX
XX Composition for prevention of alloimmunization or immunosuppression of
PT a response elicited by alloimmunization or an autoimmune hemolytic
PT disease, comprises an epitope of a rhesus protein
XX
XX Disclosure; Page 85-86; 92pp; English.
PS
XX
XX Human blood contains the rhesus (Rh) blood group system, and humans can
CC either be RhD positive or negative. This can lead to complications
CC during transfusions or pregnancy if RhD negative individuals are exposed
CC to RhD positive blood, leading to them becoming immunised to produce
CC anti-D. The present invention relates to new human allo- and
CC auto-reactive T-cell epitopes (AA99760-Y99769 and AAB03201-B03337) from
CC RhD, Rhc, Rho, Rhd and Rhe proteins. These epitopes bind to T-cells to
CC elicit an immune response, i.e. immunisation. These epitopes can be used
CC as a vaccine for the prevention of alloimmunisation or immunosuppression
CC of a response elicited by alloimmunisation or an autoimmune hemolytic
CC disease. Examples of autoimmune haemolytic diseases are sickle cell
CC disease and thalassaemia. The present sequence is the human RhCe protein
CC from which the epitopes in the present invention were derived.
XX
XX
SQ Sequence 417 AA:

Query Match 87.5%; Score 1859; DB 21; Length 417;
Best Local Similarity 89.9%; Pred. No. 6,4e-173;
Matches 375; Conservative 10; Mismatches 32; Indels 0; Gaps 0;

QY 1 MSCKPYSVGRCLPLCALTLTEALILILFEPETHYDASLEQKGVAYQVQCDDITVMAAI 60
DB 1 MSKYPRSVIRICPLCALTEALILILFETHYDASLEQKGVAYQVQCDDITVMAAI 60
QY 61 GLGFTSSFRHSSVAFLFALGVQWAILLDGFLSQPSPGKVVITLFFSILATMSA 120
DB 61 gLgftssfrhssvafllfmalgvgwailldglfsgfpgskvvtlffsrlatmsa 120
QY 121 LSVLIISDAVLGKYNLAQLVYVNLVEVTDLGNLRVVISNIFNTDYHNMMAHIYFAAYFG 180
DB 121 msvliisgavlgkynlaqlvymvivevtalgtlrmvisnifntdyhnmrlrfyfaayfg 180
QY 181 LTVAMCPRKPLREGTEHNDORATIPSLSAMIGALFLMFPSPVNSALRLRSPIERKNVFN 240
DB 181 ltvawclpkrplpygtednqratipslsamigalflwmfpvnsallrspidrknamfn 240

QY 241 TYTAAVSVVTAISGSLAHPOGKISKTYGHSALVLPDEGNAVdTSCHLIDSPWLPVLGLY 300
DB 241 tytaavsvvtaisgslahpogkisktyghsavlpgvavgschlipswlamlvlgly 300
QY 301 AGLISVYGAKYLPCCNRVLGPHSSIMGNFSLGLLEIITVILVLDTVAGAGMGIG 360
DB 301 aglissvgyaklpccnrvgllhsvmsifslglleltvylvldtvagngmgig 360
QY 361 FOYLSIGETSLAVIALTSGLTALLNKIRKAPHEAKYFPDQVFWKPHLAVGF 417
DB 361 foylslgetslavialtsgltallnkirkapheakypdqvfwkphlavgf 417

RESULT 4
AAB03338
ID AAB03338 standard; Protein; 417 AA.
AC AAB03338;
XX
XX
XX 26-SEP-2000 (first entry)
DT
XX
XX Human RhCE protein.
DE
XX
XX Human; rhesus blood group system; Rh; RhD; RhCE; sickle cell disease;
KM thalassaemia; Rhc; Rho; Rhe; Rhd; alloimmunisation prevention;
KW autoimmune Rh haemolytic disease; rhesus protein; immunosuppressive;
KW vaccine.
XX
XX Homo sapiens.
OS
XX
XX MO200032632-A2.
PN
XX
XX 08-JUN-2000.
PD
XX
XX 01-DEC-1999; 99MO-GB04027.
PE
XX
XX 01-DEC-1998; 98GB-0026378.
PR
XX
XX (UYAB-) UNIV ABERDEEN.
PA (COMM-) COMMON SERVICES AGENCY SCOTTISH HEALTH S.
XX
XX Urbanlak SJ, Barker RN;
PI
XX
XX WPI; 2000-412291/35.
DR
XX
XX Composition for prevention of alloimmunization or immunosuppression of
PT a response elicited by alloimmunization or an autoimmune hemolytic
PT disease, comprises an epitope of a rhesus protein
XX
XX Disclosure; Page 83-84; 92pp; English.
PS
XX
XX Human blood contains the rhesus (Rh) blood group system, and humans can
CC either be RhD positive or negative. This can lead to complications
CC during transfusions or pregnancy if RhD negative individuals are exposed
CC to RhD positive blood, leading to them becoming immunised to produce
CC anti-D. The present invention relates to new human allo- and
CC auto-reactive T-cell epitopes (AA99760-Y99769 and AAB03201-B03337) from
CC RhD, Rhc, Rho, Rhd and Rhe proteins. These epitopes bind to T-cells to
CC elicit an immune response, i.e. immunisation. These epitopes can be used
CC as a vaccine for the prevention of alloimmunisation or immunosuppression
CC of a response elicited by alloimmunisation or an autoimmune hemolytic
CC disease. Examples of autoimmune haemolytic diseases are sickle cell
CC disease and thalassaemia. The present sequence is the human RhCE protein
CC from which the epitopes in the present invention were derived.
XX
XX
SQ Sequence 417 AA:

Query Match 87.2%; Score 1854; DB 21; Length 417;
Best Local Similarity 89.7%; Pred. No. 2e-172;
Matches 374; Conservative 10; Mismatches 33; Indels 0; Gaps 0;

OY	1	MSKKPRSVGICLPICALTEEAAILILEFPEETHDALEDOOKGVNVSXYQGGDLTYMAAI	60
		1	
Db	1	mskyprievrriclpicaltleaaillllyfifthydaaleqdkglvasyvgqdlvmaai	60
OY	61	GIGFTSTSSFRSHSSSVAFNLFMLALGVOMAILLDGFLSQPSGKVVYTFSTLATMSA	120
		61	
Db	61	gigftstssfrshsssvafnlfmlalgvqwailldglfsgfsgkvvylcfstlatmsa	120
OY	121	LSVLISVDALGKYNLAQLVVMVLEVTDLGNLRMTSNIENFTDYHMMNMHIYFPAAYFG	180
		121	
Db	121	msvlisagavlgkynlaqlvwmvlevtalglrmvtsnifndcyhmnlrhfyvfaayfg	180
OY	181	LTVMACLRKPLPEGEEDDDQFATIPRSISAMGALFIMMFPSVNSALLRSPERRKNAVN	240
		181	
Db	181	ltvmacrpkplpkgedddqfatitpsisamgalfimfwpsvnspllrpdrknamfr	240
OY	241	TYVAAVAVVMAISGSSLAHPQGISKTYGHSAYLPQGVAVDTSCHLIPSPMPRLVILGV	300
		241	
Db	241	tyaavavvmaaisgsslahpqgisksaylpqgvavdtschlipspmprlvlgv	300
OY	301	AGLISVIGAKYLPCCCNRVLGIPHSISMGYNFSLGLLEETIYVILVLDTVGAGNGMTG	360
		301	
Db	301	aglisvgakclpcccnrvlglhishvmsifslglleetylavlhlvtwngmgig	360
OY	361	FOYLIISTEISLAIYIATISGLTALLNLKIRAPHEAKYFPDQYPMKPRHLANGF	417
		361	
Db	361	foylisigelslaivaltsglltqlllnlklwkapvakyfdqyfwkfrhlavgf	417
RESULT 5			
AAB03342			
ID	AAB03342	standard; Protein: 417 AA.	
AC	AAB03342:		
XX			
DF	26-SEP-2000	(first entry)	
XX			
DE		Human Rhce protein.	
XX			
KW	Human; rhesus blood group system; Rh; RhD; RhCE; sickle cell disease;		
KW	thalassaemia; RHC; Rhc; RHE; RHe; alloimmunisation prevention;		
KW	autoimmune Rh hemolytic disease; rhesus protein; immunosuppressive;		
KW	vaccine.		
XX			
OS	Homo sapiens.		
XX			
PN	WO200032632-A2.		
XX			
PD	08-JUN-2000.		
XX			
PF	01-DEC-1999;	99WO-GB04027.	
XX			
PR	01-DEC-1998;	98GB-0026378.	
XX			
PA	(UYAB-) UNIV ABERDEEN.		
PA	(COMM-) COMMON SERVICES AGENCY SCOTTISH HEALTH S.		
XX	Urbanlak SJ, Barker RN;		
PI			
DR	WPI: 2000-412291/35.		
XX			
XX			
PT	Composition for prevention of alloimmunization or immunosuppression of		
PT	a response elicited by alloimmunization or an autoimmune hemolytic		
PT	disease, comprises an epitope of a rhesus protein -		
CC			
CC	Disclosure; Page 90-92; 92pp: English.		
CC			
CC	Human blood contains the rhesus (Rh) blood group system, and humans can		
CC	either be RhD positive or negative. This can lead to complications		
CC	during transfusions or pregnancy if RhD negative individuals are exposed		
CC	to RhD positive blood, leading to them becoming immunised to produce		
CC	anti-D. The present invention relates to new human allo- and		
CC	auto-reactive T-cell epitopes (AA99760-999769 and AAB03201-B03337) from		

[illegible]

XX WPI; 2000-412291/35.
 XX Composition for prevention of alloimmunization or an autoimmune hemolytic
 PT response elicited by alloimmunization or an autoimmune hemolytic
 PT disease, comprises an epitope of a Rhesus protein
 XX
 PS Disclosure; Page 86-88; 92pp; English.
 CC Human blood contains the rhesus (Rh) blood group system, and humans can
 CC either be RhD positive or negative. This can lead to complications
 CC during transfusions or pregnancy if RhD negative individuals are exposed
 CC to RhD positive blood, leading to them becoming immunised to produce
 CC anti-D. The present invention relates to new human allo- and
 CC auto-reactive T-cell epitopes (AA99760-Y99769 and AAB03201-B03337) from
 CC RHD, Rhc, Rho, Rhd and Rhes proteins. These epitopes bind to T-cells to
 CC elicit an immune response, i.e. immunisation. These epitopes can be used
 CC as a vaccine for the prevention of alloimmunisation or immunosuppression
 CC of a response elicited by alloimmunisation or an autoimmune haemolytic
 CC disease. Examples of autoimmune haemolytic diseases are sickle cell
 CC disease and thalassemia. The present sequence is the human RHD protein
 CC from which the epitopes in the present invention were derived.
 CC
 XX Sequence 417 AA;
 S0
 Query Match 86.3%; Score 1833; DB 21; Length 417;
 Best Local Similarity 88.7%; Pred. No. 2.2e-170;
 Matches 370; Conservative 12; Mismatches 35; Indels 0; Gaps 0;
 QY 1 MSCKPSPVGRCLPLCALTEALILEFETRYDASLEQKGLVASYOGDITVMAAI 60
 DB 1 msckpssvrrclplwalleaiaillffethydaasleqkglvasyogdiltmaai 60
 QY 61 GLGLTSSFRHSSVAFNFMALGVQNALIDGLSPSPSKVITLFTSLATMSA 120
 DB 61 glgltsfrhssvafnfmalgvqnalidglspspskvittlftslatmsa 120
 DB 61 glgltsfrhssvafnfmalgvqnalidglspspskvittlftslatmsa 120
 QY 121 LSVTISDAVLGKYNALQVNVLEVTDLGNLBMVSNINFTYHNMHIIYFAAYFG 180
 DB 121 lsvtisdaavlgnalqvnlvlevtldgnlbnvsninfthyhnmhiiyfaayfg 180
 DB 121 msvltsgavlgkynalqvnlvlevtldgnlbnvsninfthyhnmhiiyfaayfg 180
 QY 181 LTVAMCLPKPLECTEDNDQATIPSLSAMGALFLMWFPSVSNALLSPIDERKNAFVN 240
 DB 181 ltvamclpkplctedndqatipslsamgalfmlwfpsvsnallspiderknafvn 240
 DB 181 ltvamclpkplctedndqatipslsamgalfmlwfpsvsnallspiderknafvn 240
 QY 241 TYRVANSVYTAIGSSLAHQKISTKYGSAVLPBGVAVDTCHLIPSPWLPVIGLV 300
 DB 241 tyrvansvytaigsslahqkistkygsavlpbgvavdtchlipspwlpvlgvlv 300
 DB 241 tyrvansvytaigsslahqkistkygsavlpbgvavdtchlipspwlpvlgvlv 300
 QY 301 AGLISVAGAKYLPCCNRVIGIPHSIMGVNFSLLGLEITIVLVLDVVGANGMIG 360
 DB 301 aglitsvagaklpccnrvigiphsimgvnfslgletitvllvldvvgangmig 360
 DB 301 aglitsvagaklpccnrvigiphsimgvnfslgletitvllvldvvgangmig 360
 QY 361 FOVLISIGELSLAIVALTSGLTALLNLTKIRAPHEAKYDDQVFWKPEHLAVGF 417
 DB 361 fovlisisgelslaivaltsgltallnltkirapheakyddqvfwkpehlavgf 417
 DB 361 fovlisisgelslaivaltsgltallnltkirapheakyddqvfwkpehlavgf 417

RESULT 7

AA042135
 ID AAB42135 standard; Protein: 473 AA.

XX AAB42135;

XX 08-FEB-2001 (first entry)

DE Human ORFX ORF1899 polypeptide sequence SEQ ID NO: 3798.

XX Human; open reading frame; ORFX: detection; cytostatic; hepatotropic;
 KW vulnary; antiparkinsonian; neurotropic; neuroprotective;
 KW anticonvulsant; osteoprotic; antidiabetic; immunosuppressant; cardiant;
 KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;

KW hypotensive; dermatological; immunosuppressive; antiinflammatory;
 KW antiviral; antibacterial; antifungal; antirheumatic; antihypertoid;
 KW antianemic; gene therapy; cancer; proliferative disorder; hypertension;
 KW neurodegenerative disorder; osteoarthritis; graft vs host disease;
 KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
 KW cholesterol ester storage; systemic lupus erythematosus; infection;
 KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
 KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
 KW bone damage; cartilage damage; antiinflammatory disease; coagulation;
 KW thrombosis; contraceptive.
 XX Homo sapiens.
 OS
 XX MO200058473-A2.
 PN
 XX 05-OCT-2000.
 PD
 XX 31-MAR-2000; 2000WO-0508621.
 PE
 XX 31-MAR-1999; 99US-0127607.
 PR 02-APR-1999; 99US-0127636.
 PR 05-APR-1999; 99US-0127728.
 PR 30-MAR-2000; 2000US-0540763.
 PR
 XX (CURA-) CURAGEN CORP.
 PA
 XX Shlmkets RA, Leach M;
 PI
 XX WPI; 2000-602362/57.
 DR N-PSDB; AAC76344.
 XX Novel nucleic acids and peptides derived from open reading frame X,
 PT useful for treating e.g. cancers, proliferative disorders,
 PT neurodegenerative disorders and cardiovascular disease -
 XX
 XX Claim 11; Page 2948-2949; 5507pp; English.
 XX
 CC AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
 CC which represent the human ORFX open reading frames 1 to 3161. The ORFX
 CC sequences have activities such as: cytostatic; hepatotropic; vulnary;
 CC antiparkinsonian; neurotropic; neuroprotective;
 CC osteoprotic; anticonvulsant; antidiabetic; immunosuppressant;
 CC immunostimulant; cardiant; thrombolytic; coagulant; vasotropic;
 CC antidiabetic; hypotensive; dermatological; immunosuppressive;
 CC antiinflammatory; antibacterial; antiviral; antifungal; antineumatic;
 CC antihypertoid; and antianemic. The sequences can be used for determining
 CC the presence of or predisposition to, or preventing or treating
 CC pathological conditions associated with an ORFX-associated disorder. The
 CC nucleic acids can be used to express ORFX proteins in gene therapy
 CC vectors. The proteins and nucleic acids may be used to treat cancers,
 CC proliferative disorders, neurodegenerative disorders, osteoarthritis,
 CC graft vs host disease, cardiovascular disease, diabetes mellitus,
 CC hypertension, hypothyroidism, cholesterol ester storage, systemic lupus
 CC erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,
 CC bacterial or fungal infection, malaria, autoimmune disorders, asthma,
 CC allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,
 CC nocturnal haemoglobinuria, antiinflammatory disease; to enhance
 CC coagulation; to inhibit thrombosis; and as a contraceptive.
 CC
 XX Sequence 473 AA;
 S0
 Query Match 23.4%; Score 497; DB 21; Length 473;
 Best Local Similarity 29.4%; Pred. No. 4.4e-40;
 Matches 129; Conservative 77; Mismatches 183; Indels 50; Gaps 8;
 QY 13 LPLCALTEALILEFETRYDASLEQKGLVASYOGDITVMAAI 58
 DB 5 lplcaltlwalleaiaillffethydaasleqkglvasyogdiltmaai 58
 DB 5 lplcaltlwalleaiaillffethydaasleqkglvasyogdiltmaai 58
 QY 59 AIGGLTSSFRHSSVAFNFMALGVQNALIDGLSPSPSKVITLFTSLATMSA 118
 DB 65 aiggltsfrhssvafnfmalgvqnalidglspspskvittlftslatmsa 118
 DB 65 aiggltsfrhssvafnfmalgvqnalidglspspskvittlftslatmsa 118

PR 25-MAY-1999; 9905-0136021.
PR 27-MAY-1999; 9905-0136392.
PR 28-MAY-1999; 9905-0136782.
PR 01-JUN-1999; 9905-0137222.
PR 03-JUN-1999; 9905-0137528.
PR 04-JUN-1999; 9905-0137502.
PR 07-JUN-1999; 9905-0137724.
PR 08-JUN-1999; 9905-0138094.
PR 10-JUN-1999; 9905-0138540.
PR 10-JUN-1999; 9905-0138847.
PR 14-JUN-1999; 9905-0139119.
PR 16-JUN-1999; 9905-0139452.
PR 16-JUN-1999; 9905-0139453.
PR 17-JUN-1999; 9905-0139492.
PR 18-JUN-1999; 9905-0139454.
PR 18-JUN-1999; 9905-0139455.
PR 18-JUN-1999; 9905-0139456.
PR 18-JUN-1999; 9905-0139457.
PR 18-JUN-1999; 9905-0139458.
PR 18-JUN-1999; 9905-0139459.
PR 18-JUN-1999; 9905-0139460.
PR 18-JUN-1999; 9905-0139461.
PR 18-JUN-1999; 9905-0139462.
PR 18-JUN-1999; 9905-0139463.
PR 18-JUN-1999; 9905-0139750.
PR 18-JUN-1999; 9905-0139763.
PR 21-JUN-1999; 9905-0139817.
PR 22-JUN-1999; 9905-0140353.
PR 23-JUN-1999; 9905-0140354.
PR 23-JUN-1999; 9905-0140354.
PR 24-JUN-1999; 9905-0140695.
PR 28-JUN-1999; 9905-0140823.
PR 29-JUN-1999; 9905-0140991.
PR 30-JUN-1999; 9905-0141287.
PR 01-JUL-1999; 9905-0141842.
PR 01-JUL-1999; 9905-0142154.
PR 02-JUL-1999; 9905-0142055.
PR 06-JUL-1999; 9905-0142380.
PR 08-JUL-1999; 9905-0142803.
PR 09-JUL-1999; 9905-0142920.
PR 12-JUL-1999; 9905-0142977.
PR 13-JUL-1999; 9905-0143542.
PR 14-JUL-1999; 9905-0143624.
PR 15-JUL-1999; 9905-0144005.
PR 16-JUL-1999; 9905-0144085.
PR 16-JUL-1999; 9905-0144086.
PR 19-JUL-1999; 9905-0144325.
PR 19-JUL-1999; 9905-0144331.
PR 19-JUL-1999; 9905-0144332.
PR 19-JUL-1999; 9905-0144333.
PR 19-JUL-1999; 9905-0144334.
PR 19-JUL-1999; 9905-0144334.
PR 19-JUL-1999; 9905-0144335.
PR 20-JUL-1999; 9905-0144352.
PR 20-JUL-1999; 9905-0144632.
PR 20-JUL-1999; 9905-0144844.
PR 21-JUL-1999; 9905-0144814.
PR 21-JUL-1999; 9905-0145086.
PR 21-JUL-1999; 9905-0145086.
PR 22-JUL-1999; 9905-0145085.
PR 22-JUL-1999; 9905-0145087.
PR 22-JUL-1999; 9905-0145089.
PR 22-JUL-1999; 9905-0145089.
PR 22-JUL-1999; 9905-0145192.
PR 23-JUL-1999; 9905-0145145.
PR 23-JUL-1999; 9905-0145218.
PR 23-JUL-1999; 9905-0145224.
PR 26-JUL-1999; 9905-0145276.
PR 27-JUL-1999; 9905-0145913.
PR 27-JUL-1999; 9905-0145918.
PR 27-JUL-1999; 9905-0145919.
PR 28-JUL-1999; 9905-0145951.
PR 02-AUG-1999; 9905-0146386.
PR 02-AUG-1999; 9905-0146388.
PR 02-AUG-1999; 9905-0146389.

PR 03-AUG-1999; 9905-0147038.
PR 04-AUG-1999; 9905-0147204.
PR 04-AUG-1999; 9905-0147302.
PR 05-AUG-1999; 9905-0147192.
PR 05-AUG-1999; 9905-0147260.
PR 06-AUG-1999; 9905-0147303.
PR 06-AUG-1999; 9905-0147416.
PR 09-AUG-1999; 9905-0147493.
PR 09-AUG-1999; 9905-0147935.
PR 10-AUG-1999; 9905-0148171.
PR 11-AUG-1999; 9905-0148319.
PR 12-AUG-1999; 9905-0148341.
PR 13-AUG-1999; 9905-0148365.
PR 13-AUG-1999; 9905-0148684.
PR 16-AUG-1999; 9905-0149368.
PR 17-AUG-1999; 9905-0149175.
PR 18-AUG-1999; 9905-0149426.
PR 20-AUG-1999; 9905-0149722.
PR 20-AUG-1999; 9905-0149723.
PR 20-AUG-1999; 9905-0149929.
PR 23-AUG-1999; 9905-0149902.
PR 23-AUG-1999; 9905-0149930.
PR 25-AUG-1999; 9905-0150566.
PR 26-AUG-1999; 9905-0150884.
PR 27-AUG-1999; 9905-0151065.
PR 27-AUG-1999; 9905-0151066.
PR 27-AUG-1999; 9905-0151080.
PR 30-AUG-1999; 9905-0151303.
PR 31-AUG-1999; 9905-015138.
PR 01-SEP-1999; 9905-0151330.
PR 07-SEP-1999; 9905-0152363.
PR 10-SEP-1999; 9905-0153070.
PR 13-SEP-1999; 9905-0153758.
PR 15-SEP-1999; 9905-0154018.
PR 16-SEP-1999; 9905-0154039.
PR 20-SEP-1999; 9905-0154779.
PR 22-SEP-1999; 9905-0155139.
PR 23-SEP-1999; 9905-0155139.
PR 24-SEP-1999; 9905-0155486.
PR 28-SEP-1999; 9905-0155659.
PR 29-SEP-1999; 9905-0156458.
PR 04-OCT-1999; 9905-0156596.
PR 04-OCT-1999; 9905-0157117.
PR 05-OCT-1999; 9905-0157753.
PR 06-OCT-1999; 9905-0157865.
PR 07-OCT-1999; 9905-0158029.
PR 08-OCT-1999; 9905-0158232.
PR 12-OCT-1999; 9905-0158369.
PR 13-OCT-1999; 9905-0159293.
PR 13-OCT-1999; 9905-0159294.
PR 13-OCT-1999; 9905-0159295.
PR 14-OCT-1999; 9905-0159329.
PR 14-OCT-1999; 9905-0159330.
PR 14-OCT-1999; 9905-0159331.
PR 14-OCT-1999; 9905-0159637.
PR 18-OCT-1999; 9905-0159638.
PR 18-OCT-1999; 9905-0159584.
PR 21-OCT-1999; 9905-0160741.
PR 21-OCT-1999; 9905-0160767.
PR 21-OCT-1999; 9905-0160768.
PR 21-OCT-1999; 9905-0160770.
PR 21-OCT-1999; 9905-0160814.
PR 21-OCT-1999; 9905-0160815.
PR 22-OCT-1999; 9905-0160980.
PR 22-OCT-1999; 9905-0160981.
PR 22-OCT-1999; 9905-0160989.
PR 25-OCT-1999; 9905-0161404.
PR 25-OCT-1999; 9905-0161405.
PR 25-OCT-1999; 9905-0161406.
PR 26-OCT-1999; 9905-0161359.
PR 26-OCT-1999; 9905-0161360.
PR 26-OCT-1999; 9905-0161361.
PR 28-OCT-1999; 9905-0161920.
PR 28-OCT-1999; 9905-0161992.

AAW98770;
 31-MAR-1999 (first entry)
 H. pylori GHPO 1130 protein.
 GHPO protein; Helicobacter infection; gastroduodenal disease; gastritis;
 peptic ulcer disease.
 Helicobacter pylori.
 WO9843478-A1.
 08-OCT-1998.
 01-APR-1998; 98WO-US06371.
 29-JUL-1997; 97US-0902615.
 01-APR-1997; 97US-0833457.
 24-JUN-1997; 97US-0881227.
 (HUMA-) HUMAN GENOME SCI INC.
 (INMR) MERIEUX ORAVAX PASTEUR MERIEUX SERUMS.
 Al-Garawi A, Kleantous H, Miller C, Comen RP, Tomb J;
 WPI: 1998-542293/46.
 N-PSDB; AAX144489.
 New isolated Helicobacter polynucleotides - used to develop products
 for the diagnosis, prevention and treatment of Helicobacter
 infections and gastrointestinal diseases
 Claim 8; Page 1663-1665; 2054pp; English.
 This sequence represents a Helicobacter pylori GHPO protein of the
 invention. The polypeptides can be used for preventing or treating
 Helicobacter infections, and gastroduodenal diseases associated with
 these infections, including acute, chronic, and atrophic gastritis, and
 peptic ulcer diseases, e.g. gastric and duodenal ulcers. They can also be
 used for the production of antibodies. The products can also be used for
 detection and diagnosis.

Sequence 449 AA;

Query Match 4.9%; Score 104.5; DB 19; Length 449;
 Best Local Similarity 20.8%; Pred. No. 0.076;
 Matches 85; Conservative 63; Mismatches 114; Indels 147; Gaps 20;

27 LFFEFTHYDASLEDDOKGLVASYQVGDLTVMMAIGLFTSSFRHSSVAFNLEMLAL 86
 130 llyff--lhnatpugyf-----glpmatclafalgvmllygr--vptalkvlltcl 178
 87 GYVMAIILDDGLSOPSPGKVVITLF-----SIWATMSALSVLISDAVAGKVNLAOLV 140
 179 av-----addlgavlvialfytnlkfawl--lgalgvvl-vlavlnlmmsl 225
 141 VMVLVEVTDGNNRMWYSINFTNDYHMMNMHTYVFAAYFGLVAMCLPKPLPESTED--- 197
 226 pylllygv-----llwfcvbsgihataa---vllaflmipvkfipkdsknvel 269
 198 -----ndoratiPSLSAMLGALFLMMP-PSVNSALRSPIERKNVFN 240
 270 lclgkyraetsgallskqegelhlsleekasa-----lqspier----- 309
 241 TYVAAVSVVTAISGSLAHPOGRISKITYGSAVLP-----EGVAVDTSCHLIPSEWLP 294
 310 -----lehlfiaplsygf-----lmpifafanagvsdssinlevd---k 345
 295 lVIGIVAGLISVCAKYLPGCCNRVIGIRHSSIMGYNFSLGLLAEITTYVLVLDIVGA 354
 346 vllgvllyl-----clgkplgylflitfiseklklitarpkgiswvhl-----ga 389

0Y 355 G-NGMICEVOLLSTIGELS-----LAIVL-ALTSGLLTALL 388
 390 gllagjgftmsmflsnlatfscnkhdamewakiallglslisgllgalyl 438

RESULT 13
 AAW20821
 ID AAW20821 standard; Protein; 497 AA.
 XX
 AC AAW20821;
 16-JUL-1997 (first entry)
 H. pylori cell envelope inner membrane protein 11cel1603orf16.
 Cytoplasmic; vaccine; prevention; treatment; infection; identification;
 binding compound; bacterium; life cycle; activator; bacteria; inhibitor;
 duodenal ulcer disease; chronic gastritis; diagnosis; envelope.
 Helicobacter pylori.
 WO9640893-A1.
 19-DEC-1996.
 06-JUN-1996; 96WO-US09122.
 01-APR-1996; 96US-0630405.
 07-JUN-1995; 95US-0487032.
 (ASTR) ASTRA AB.
 Berglundh OF, Smith D, Mellgaerd BL;
 WPI: 1997-052306/05.
 N-PSDB; AAT68074.
 Helicobacter pylori nucleic acid sequences and related
 polypeptide(s) - useful for vaccines to treat or prevent H. pylori
 infection, and to detect Helicobacter
 Claim 56; Page 1226-1227; 1481pp; English.

The present sequence shows a Helicobacter pylori cell envelope
 inner membrane protein.
 The protein may be used in a vaccine to prevent or treat H. pylori
 infection or to identify H. pylori polypeptide binding compounds,
 useful as potential H. pylori life cycle activators or inhibitors.
 The genomic sequence of H. pylori (ATCC 55679) was determined from
 overlapping contigs generated by mechanically shearing the bacterial
 DNA. The sequences were analysed for ORF of at least 180 nucleotides,
 and the predicted coding regions defined by computer evaluation. To
 identify likely H. pylori antigens for vaccine development, the amino
 acid sequences predicted from various ORF were analysed for significant
 homology to other known or exported membrane proteins. Having identified
 and determined the sequences of interest, particular regions can be
 isolated from H. pylori by PCR amplification for recombinant polypeptide
 production, e.g. in E. coli hosts.

Sequence 497 AA;

Query Match 4.9%; Score 104.5; DB 18; Length 497;
 Best Local Similarity 22.7%; Pred. No. 0.086;
 Matches 111; Conservative 52; Mismatches 174; Indels 153; Gaps 27;

15 LCAUTTEAAILLFFFTTH---YDASLEDDOKGLVASYQVGDLTVMMAI----- 60
 48 lcvlvsfigavvllfigqwhgsyekylfdwi-vvgntkxgfsImldnlnvmlvvlvs 106
 61 -----GLGFLV--TSSRRHSSVAFNLMALGVOMAILDDGLSOPSPS-----GKVV 108

```

Db 107 flvhvyslygmehdgtfnryfyslsgfismlyl-----vlsdnflglfigwewgvlcsy 161
QY 109 TLFSTVATMSALSVLISVDVAGVKVNLADLVVVLVEVTDLGNL----- 153
Db 162 lllgfwyhhksa-----nnasleatvmrirdlglmlnglllfnfgclgy 207
QY 154 RMVTSNFTNDYHNMNMHIYF-----AAVFGT--LYAWCLPKPLPBGTEDDNDQRAF 203
Db 208 kevfsmInnadysm-lfyisvflfigmgksqgfmhtwlanamegctp----- 255
QY 204 IPSLSAALGLFLMF--FPVNSALRSPIERKNVNT--YAAVAVSVVTATSGSLAHPQ 262
Db 256 ---vsallth--atmvtaagyvlltrnpl--ysavfevgyfiacigafvalfgasma1vn 308
QY 263 GKISRTYGHSAVLPEG---VAVDTSC-----HLIPSPMPrlVGLVAGL----- 303
Db 309 kdlkrtvaystlsglgymfraaglgayalalfnhftchafksllflfsgnmvhamednd 368
QY 304 ISVYGAAYLPDCCNRVLGPHSSIM-----GYNFS-----LGLLGEIYIVLLVL 349
Db 369 ltkmgalykpmrltavfmligsvalcglypfagy-fskdklleafgmbhhlwfvll-- 425
QY 350 DYVAGNGMIGFQVLLSIGELSLAIVATLSGLTALLMLKIRKAPHEAKY-FDDQYFW 408
Db 426 -----lg-----alfcayfsfllmlvf---fapkdheinhphgqkfh 461
QY 409 KFPPLAV-GF 417
Db 462 afehatvrgf 471

```

RESULT 14

AA81556
ID AAY81556 standard; Protein: 449 AA.

XX AAY81556;

XX 24-MAY-2000 (first entry)

XX Streptococcus pneumoniae type 4 protein sequence #56.

XX Streptococcus pneumoniae; vaccine; screening; protein antigen.

XX Streptococcus pneumoniae; vaccine; screening; protein antigen.

XX Streptococcus pneumoniae; vaccine; screening; protein antigen.

XX Streptococcus pneumoniae; vaccine; screening; protein antigen.

XX Streptococcus pneumoniae; vaccine; screening; protein antigen.

XX Streptococcus pneumoniae; vaccine; screening; protein antigen.

XX Streptococcus pneumoniae; vaccine; screening; protein antigen.

XX Streptococcus pneumoniae; vaccine; screening; protein antigen.

XX Streptococcus pneumoniae; vaccine; screening; protein antigen.

XX Streptococcus pneumoniae; vaccine; screening; protein antigen.

XX Streptococcus pneumoniae; vaccine; screening; protein antigen.

XX Streptococcus pneumoniae; vaccine; screening; protein antigen.

XX Streptococcus pneumoniae; vaccine; screening; protein antigen.

XX Streptococcus pneumoniae; vaccine; screening; protein antigen.

XX Streptococcus pneumoniae; vaccine; screening; protein antigen.

XX Streptococcus pneumoniae; vaccine; screening; protein antigen.

XX Streptococcus pneumoniae; vaccine; screening; protein antigen.

CC and/or antigens. The nucleotide sequences can be used in vaccines and in
CC diagnostic assays. The proteins and nucleotides can be useful for the
CC detection and diagnosis of S. pneumoniae. The protein sequences are also
CC useful for screening an agent capable of antagonizing, inhibiting or
CC interfering with the function or expression of the proteins in which the
CC agent is useful for treatment or prophylaxis of S. pneumoniae infection
CC and meningitis. AA05591 to AA05614 represent primers used in the
CC exemplification of the present invention.

SO Sequence 449 AA:

Query Match 4.9%; Score 104; DB 21; Length 449;
Best Local Similarity 22.0%; Pred. No. 0.085;
Matches 85; Conservative 53; Mismatches 149; Indels 100; Gaps 17;

```

QY 43 GLVASVOVGDLV--MAALGLGFLTSSFRHMSVAFLFALALGVCAIILLDGLSQ 100
Db 20 gktaqfkvvraitaagaaavflcivgsmflvslpqaif-----wplvadlfsas 71
QY 101 FPGSKVYITLFSIWL-----ATMSALSVLISVDVAGVKVNLADLVVVLVEVTDLGNLRMV 156
Db 72 fdk-----ftslymwanyatmgsislyf-----vlsayelktlyeaeel----- 112
QY 157 ISNIFNTDYHNMNMHIYFPAAY-FGLTVAMCLPKPLPBGTEDDNDQAFATPSLSAMCALP 215
Db 113 -----nmprlngallalmafvmfv-----pqllfdggnmkvtvtslke--gav1 153
QY 216 L--WMF--PSVNSALRSPIERKNVNT--YAAVAVSVVTATSGSLAHPQKISTY 270
Db 154 adgwangnv-----arfgtclgltfcalimalavtlvlymcvkhwvlm-- 197
QY 271 HSAVLPEGVAVDTSCNLPSPMLrlVGLVAGLISVYGAAYLPDCCNRVLGPHSSIMGY 330
Db 198 -peavpegvargfta-lvpgfvavafvifingllvamgtlif-----kvalafgfvsn1 250
QY 331 NFSLLGLLEIYIVLLVLDFTVAGANGMIGFQVLLSI-----GELSLA 373
Db 251 tnswigim--llylltqlwlgvlganlvfafvspialamaanaaghfavagetsm 308
QY 374 IVATLSGLTALLMLKIRKAPHEAK 400
Db 309 fvaagsgac1g1c1yafaskseq1k 335

```

RESULT 15

AA878999
ID AAB78999 standard; Protein: 465 AA.

XX AAB78999;

XX 30-APR-2001 (first entry)

XX C. glutamicum SRT protein sequence SEQ ID NO:258.

XX Corynebacterium glutamicum; stress; resistance; tolerance; SRT;

XX fine chemical production; organic acid; pyrimidine base; nucleoside;

XX nonproteinogenic amino acid; purine base; pyrimidine base; nucleoside;

XX nucleotide; lipid; saturated fatty acid; unsaturated fatty acid; diol;

XX carbohydrate; aromatic compound; vitamin; cofactor; polyketide; enzyme;

XX evolutionary study; environmental hazard; fermentation.

XX Corynebacterium glutamicum.

XX WO200100804-A2.

XX 04-JAN-2001.

XX 23-JUN-2000; 2000MO-IB00922.

XX 25-JUN-1999; 99US-0141031.

XX 01-JUL-1999; 99DE-1030429.

XX 01-JUL-1999; 99US-0142692.

PR 08-JUL-1999; 99DE-1031413.
PR 08-JUL-1999; 99DE-1031457.
PR 08-JUL-1999; 99DE-1031541.
PR 09-JUL-1999; 99DE-1032209.
PR 09-JUL-1999; 99DE-1032230.
PR 14-JUL-1999; 99DE-1032914.
PR 27-AUG-1999; 99DE-1040764.
PR 27-AUG-1999; 99US-0151214.
PR 31-AUG-1999; 99DE-1041382.
XX
XX (BADI) BASF AG.
XX
XX Pompejus M, Kroegeer B, Schroeder H, Zelder O, Habenerauer G, Lee H;
PI Kim H;
XX
XX MPI: 2001-061972/07.
DR N-PSDB; AAF71112.
XX
XX

PT New isolated Corynebacterium glutamicum nucleic acid encoding a stress,
PT tolerance or resistance protein, for production or modulation of
PT production of fine chemicals, such as, e.g. amino acids, lipids,
PT carbohydrates, or enzymes -
XX
XX

PS Claim 20; Page 465-466; 526pp; English.

XX AAF70984 to AAF71133 encode the Corynebacterium glutamicum stress,
CC resistance, and tolerance (SRT) proteins given in AAB78871 to AAB79020.
CC The C. glutamicum SRT genes (I) can be used in vectors (II) for
CC expression in host cells and production of fine chemicals, such as, an
CC organic acid, a proteinogenic or nonproteinogenic amino acid (preferred),
CC a purine or pyrimidine base, a nucleoside, a nucleotide, a lipid, a
CC saturated or unsaturated fatty acid, a diol, a carbohydrate, an aromatic
CC compound, a vitamin, a cofactor, a polyketide, or an enzyme. The SRT
CC chemical production can be modulated. The presence of (I) or the SRT
CC proteins (III) encoded by them are used for diagnosing the presence
CC or activity of Corynebacterium diptheriae. (I), (II), (III) and host
CC cells containing them can be used to map the genomes of organisms related
CC to C. glutamicum, to identify and localise C. glutamicum sequences of
CC interest, in evolutionary studies, in determination of SRT protein
CC regions required for function, in modulating the SRT protein activity,
CC and in modulating the activity of an SRT pathway. (II) are used to permit
CC C. glutamicum to survive in an environment that is normally
CC environmentally or chemically hazardous to it. (I) and protein molecules
CC encoded by it increase the survival of C. glutamicum to chemical and
CC environmental hazards and provide a means for continued growth and
CC multiplication in large scale fermentative growth conditions. By
CC increasing the growth rate or maintaining a normal growth rate in poor or
CC toxic conditions, the yield, production and/or efficiency or production
CC of fine chemicals from a culture may be increased.
XX
XX

SQ Sequence 465 AA:

Query Match 4.9%; Score 104; DB 22; Length 465;
Best Local Similarity 19.8%; Pred. No. 0.089;
Matches 79; Conservative 64; Mismatches 134; Indels 122; Gaps 17;

OY 13 LPTCALTL-----EAILILFFFTHYDASLEDOKGLVASYQVGDLTVMATGLGF 64
DB 104 vpicllllrqvnekyall-----gltvsvng--lgyvdalagw 145
OY 65 L--TSSFRHSMSVAFNLGLGVOMATILDOFLSOPPSGKVITLFSIWLATMSALS 122
DB 146 laetlgrfslfwmaaf-cavaalalpfsv-keetaeepk-----mdwl---gvlp 192
OY 123 VLISVDATL-----SKVNLAQLVVNVLVVETDIGNLRVVISNIFNTDYHNMNHIVFA 176
DB 193 lavisgslmafeagllgaanwllvvllffligagv-----lffynlekrvkhplsv 246
OY 177 AYFGCLTAVKCLPRPLPGCTEDNDQRTATIPSSAMLGALFLMFPSPVNSALRLSPIERKN 236
DB 247 eylgqrtrwail-----stlltmtgvfawmgllpnlagdaanga----- 287

OY 237 AVENTYYAVASVVT-----AISGSSLAHPQGIKSTYGHSAVLPEGVAVDTSCHLIP 289
DB 288 -----gmsasvswwtlltpyalaglvfsgplagllagrkfykvlvq----- 327
OY 290 SPWLPIVLGLVAGLISVRCAGKYLPGCCNRY-----LGIPHSSTMGVNFSLGLLEET 341
DB 328 -----lgaactllgvagatflvgstshlaylglstfygltvgyanlmnglg----- 375
OY 342 IYIVLLVLDTVGAGNGMIGFOVLSTIGELSLATVIALTS 380
DB 376 --lvlsppanngylpygmamagafmaga-fsfailfavst 411

Search completed: September 12, 2001, 07:38:44
Job time: 329 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OW protein - protein search, using sw model

Run on: September 12, 2001, 07:33:51 : Search time 12.41 Seconds
(without alignments)
693.535 Million cell updates/sec

Title: 09-600714-2
Perfect score: 2125
Sequence: 1 MSCXPRSVGRCLPLCALTL.....AKYPDQVFMKFPHLAVGFX 418

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 197339 seqs, 20590346 residues

Total number of hits satisfying chosen parameters: 197339

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/prodata/1/1aa/5A_COMB.pep:*
2: /cgn2_6/prodata/1/1aa/5B_COMB.pep:*
3: /cgn2_6/prodata/1/1aa/6A_COMB.pep:*
4: /cgn2_6/prodata/1/1aa/6B_COMB.pep:*
5: /cgn2_6/prodata/1/1aa/PCTUS_COMB.pep:*
6: /cgn2_6/prodata/1/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1833	86.3	417	1 US-08-553-888A-3	Sequence 3, Appli
2	98	4.6	462	2 US-08-898-976-2	Sequence 2, Appli
3	98	4.6	462	2 US-08-898-976-4	Sequence 4, Appli
4	93.5	4.4	746	2 US-08-785-431-4	Sequence 2, Appli
5	93.5	4.4	788	2 US-08-785-431-2	Sequence 2, Appli
6	90.5	4.3	1248	2 US-08-726-214-16	Sequence 16, Appli
7	90.5	4.3	509	2 US-09-031-392-6	Sequence 6, Appli
8	90.5	4.3	509	4 US-09-299-549-6	Sequence 6, Appli
9	90	4.2	574	4 US-09-383-586-36	Sequence 36, Appli
10	90	4.2	759	4 US-08-637-759B-89	Sequence 89, Appli
11	90	4.2	759	3 US-08-871-355A-89	Sequence 89, Appli
12	88	4.1	357	1 US-08-119-773-2	Sequence 2, Appli
13	87	4.1	357	1 US-08-119-773-4	Sequence 4, Appli
14	87	4.1	357	1 US-08-119-773-6	Sequence 6, Appli
15	85.5	4.0	603	4 US-09-097-889-23	Sequence 23, Appli
16	85	4.0	336	4 US-08-749-816-2	Sequence 2, Appli
17	85	4.0	457	2 US-08-882-704A-6	Sequence 6, Appli
18	84.5	4.0	348	1 US-08-176-126B-2	Sequence 2, Appli
19	84.5	4.0	348	2 US-08-669-435-2	Sequence 2, Appli
20	84.5	4.0	348	5 PCT-US94-14431A-2	Sequence 2, Appli
21	84.5	4.0	473	1 US-08-597-236-13	Sequence 13, Appli
22	84.5	4.0	473	1 US-08-746-682A-13	Sequence 13, Appli
23	84	4.0	311	2 US-08-794-216-1	Sequence 1, Appli
24	84	4.0	429	2 US-08-677-049-5	Sequence 5, Appli
25	83.5	3.9	503	1 US-08-464-840-4	Sequence 4, Appli
26	83.5	3.9	503	1 US-08-483-094-4	Sequence 4, Appli
27	83.5	3.9	648	4 US-08-800-291B-8	Sequence 8, Appli

28	82.5	3.9	397	1 US-08-476-000-63	Sequence 63, Appli
29	82.5	3.9	397	1 US-08-472-840-63	Sequence 63, Appli
30	82.5	3.9	397	2 US-08-476-976-63	Sequence 63, Appli
31	82.5	3.9	397	3 US-08-474-410-63	Sequence 63, Appli
32	82	3.9	462	3 US-08-788-231A-15	Sequence 15, Appli
33	82	3.9	588	2 US-09-005-232A-2	Sequence 2, Appli
34	81.5	3.8	525	1 US-08-356-340-2	Sequence 2, Appli
35	81.5	3.8	525	1 US-08-786-555-2	Sequence 2, Appli
36	81	3.8	1285	2 US-08-540-406-6	Sequence 6, Appli
37	81	3.8	1285	4 US-08-656-055-6	Sequence 6, Appli
38	81	3.8	1285	4 US-08-954-668-6	Sequence 6, Appli
39	81	3.8	1285	5 PCT-US95-13233-6	Sequence 6, Appli
40	81	3.8	1286	4 US-09-268-140-3	Sequence 5, Appli
41	80.5	3.8	607	4 US-08-969-683A-59	Sequence 59, Appli
42	80.5	3.8	764	1 US-08-424-567-2	Sequence 2, Appli
43	80.5	3.8	764	2 US-08-711-928-2	Sequence 2, Appli
44	80.5	3.8	764	4 US-09-184-937-2	Sequence 2, Appli
45	80	3.8	999	2 US-08-473-533A-5	Sequence 5, Appli

ALIGNMENTS

```
RESULT 1
US-08-553-888A-3
: Sequence 3, Application US/08553888A
: Patent No. 5723293
: GENERAL INFORMATION:
: APPLICANT: Huang
: TITLE OF INVENTION: DIAGNOSTIC METHOD AND KIT FOR
: TITLE OF INVENTION: DETERMINING RH BLOOD GROUP GENOTYPE
: NUMBER OF SEQUENCES: 3
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Hoffmann & Baron, LLP
: STREET: 350 Jericho Turnpike
: CITY: Jericho
: STATE: New York
: COUNTRY: USA
: ZIP: 11753
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage
: COMPUTER: IBM compatible
: OPERATING SYSTEM: MS-DOS
: SOFTWARE: WordPerfect
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/553,888A
: FILING DATE: 11/06/95
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: O'Dea, Sean W.
: REGISTRATION NUMBER: 37690
: REFERENCE/DOCKET NUMBER: 454-5
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (516) 822-3550
: TELEFAX: (516) 822-3582
: INFORMATION FOR SEQ ID NO: 3:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 417 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: peptide
: US-08-553-888A-3

Query Match      86.3%      Score 1833; DB 1; Length 417;
Best Local Similarity 88.7%      Pred. No. 9.8e-185;
Matches 370; Conservative 12; Mismatches 35; Indels 0; Gaps 0;

QY 1 MSCXPRSVGRCLPLCALTLLEFPTHTYDASLEDDKGLVASYOGQDTLVAAI 60
   |||.....|.....|.....|.....|.....|.....|.....|.....|
Db 1 MSSXPRSVGRCLPLCALTLLEFPTHTYDASLEDDKGLVASYOGQDTLVAAL 60
QY 61 GLGFLTSSFRHRSSVAFNLFMLALGYOMAILLDGLFSOPPSGKVYITLFSIWLATMSA 120
```

Db	61	GLGFLTISNFRHSHMSVAFNLFMLALGYQAAIILDLGFLSOPPECKVYITTFPIRLATMSA	1200
Qy	121	LSVLISVDALIGKYNLAQLVYVYVLEVTDLGLENRMVTSINFTNDYHNMKHIIVEAAFG	1800
Db	121	MSVLISAGAVLGKYNLAQLVYVYVLEVTALGTLERMVTSINFTNDYHNMLEHFEVFAAFC	1800
Qy	181	LTVMACLPKPLPEGEDDDQCATIPISLSAMIGALFEMMFPSVSAALRSPIERKNAFEN	2400
Db	181	LTVAMCCLPKPLPKGEDDDQCATIPISLSAMIGALFEMMFPSVSAALRSPIERKNAFEN	2400
Qy	241	TYVAANVSVAALIGSSSLAHPOGRISKTYGSHAVLPGEVAVGTSCHLIPSPMLIPYLGIV	3000
Db	241	TYVAALAVSVVAALIGSSSLAHPOGRISIMTYVHSAVLAGGAVAVGTSCHLIPSPMLAVLGIV	3000
Qy	301	AGLISVIGAKTLPGCCCNFVLGIPHSITMGVNFSLGLLEETIYVLLVDTPVAGNGMIG	3600
Db	301	AGLISVIGAKCLPYCCCNFVLGIHHSIVMHSIFSLGLLEGTIYVLLVLTGVNNGMIG	3600
Qy	361	FOVLIISTELSLAVIALTSLGLLALLNLKIFAPRHAAYEPDQVQWKRPHLAUGF	417
Db	361	FOVLIISTELSLAVIALTSLGLLNLKIKAPRHAAYEPDQVQWKRPHLAUGF	417

RESULT 2
 US-08-898-976-2
 Sequence 2, Application US/08898976
 Patent No. 5891670
 GENERAL INFORMATION:
 APPLICANT: Burnham, Martin
 APPLICANT: Lonetto, Michael
 APPLICANT: Warren, Patrick
 TITLE OF INVENTION: NOVEL TETRACYCLINE RESISTANCE
 TITLE OF INVENTION: PROTEIN
 NUMBER OF SEQUENCES: 6
 CORRESPONDENCE ADDRESSES:
 ADDRESSEE: Dechart Price & Rhoads
 STREET: 997 Lenox Drive, Building 3, Suite 210
 CITY: Lawrenceville
 STATE: NJ
 COUNTRY: USA
 ZIP: 08543
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: DOS
 SOFTWARE: FastSeq for Windows Version 2.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/898,976
 FILING DATE:
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER:
 FILING DATE:
 ATTORNEY/AGENT INFORMATION:
 NAME: Bloom, Allen
 REGISTRATION NUMBER: 29,135
 REFERENCE/DOCKET NUMBER: GM10044
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 609-520-3214
 TELEFAX: 609-520-3259
 TELEX:
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 462 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 US-08-898-976-2

Query Match	4.68;	Score 98;	DB 2;	Length 462;
Best Local Similarity	20.08;	Pred. NO.	0.065;	

	Matches	96;	Conservative	70;	Mismatches	162;	Indels	152;	Gaps	21;
QY	1	MSCKPRSVQRCPLCALTEALAILLFYFFTHNDASLEDQKGLVASVQVGODLTVMMAI	60							
Db	13	MNVEVK-IKKRVPI-----LLEFLFVSLVINDSEK-----LISVAIDDLMI-----	54							
QY	61	GLGFLTSFRRHSMSVAFNFMALGQOMAILDGLFSQ--FPESKVIVLFESLWATM	118							
Db	55	-----SVTVYSQWATLGLAVIGKGAIVYASLSAISIRPFITGVILLIIFGSLIGIF	107							
QY	119	SALSVLISVDVAGKVNLAQLVVMVLEV-----TdlG-----NLRMVISN--	159							
Db	108	QHOSPLTTEVGRITQAGALAAETLVIIVAAVYLSKEDQKTLGISTSSYSLSLVIGTLG	167							
QY	160	--IENFDVHMNMNHLYFAAFGLGVAMCIP--KRLPEGEDDQCRITPISLSMICAL	214							
Db	168	GFIFLYLHMTNELL-----ALIVYETLPLEFLKLR--ENNTKHHLDVGLILVAT	218							
QY	215	-----FLMFRPVSALL-----RSPIERKNAVFNTYAAVAVSVTAIS	254							
Db	219	IATVWLFITFENNMVLIGALIAITVFAALYTKNAQRPLVNSSEFQNKRYASFLTFVWY	278							
QY	255	GSSLHHPGQKISKITYGHSNVP-----EGVAVG-TSCHLIPSPRLPIVLGIAGLSIVR	307							
Db	279	AIQL-----GYLTFPPEIMEQIYHQLODPTTSLLPVGYVAIVAGLSKTI---	324							
QY	308	GAKYLPGCCNRYLGIPIHSSIMGYNSLGLLEETIYIV-LVLDVTGAGNGMIGEOVL	365							
Db	325	-GEYLN-----SKQAITAILLALSILLPFAFAGNHHSIFVISMI	364							
QY	366	-----SIGELSLAIYALTSGL-LtallLNLRKIRAP	396							
Db	365	FFAGSFALMVAPLNEAIKTDLNMNTGVAIGYNNLIVAAVSGVIAAALIDFALNRP	424							

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525

```

; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 462 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-898-976-4

```

```

Query Match          4.6%; Score 98; DB 2; Length 462;
Best Local Similarity 20.0%; Pred. No. 0.065;
Matches 96; Conservative 70; Mismatches 162; Indels 152; Gaps 21;

```

```

QY 1 MSCKPRSVGRCLPLCALTEALILLEYFETHYDASLEDOGLVASQVODTVMAAI 60
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 13 MNVEYSK-IKAVPI-----LLFLFVSLVIDNSFK-----LISVALADLNI----- 54
QY 61 GLGFLTSSFRHSSVAFNLFMLALGVOMAILDGLSQ--PPSGKAVITLESILATM 118
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 55 -----SVTVSWQATLAGLVIGMGAVVVASLSDAISIRPPFIYGVILLIFGSIIGNIF 107
QY 119 SALSVLISVAVLGKVMIAQVNVVAVEV-----TDLG-----NLRMTISN--- 159
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 108 QHOSPLVLGRITDTAGLAAETLYIVYAKYLSKEDQKTYLGLSTSSYSLSLVTLSG 167
QY 160 --FENEDYHNMNMHIYVFAAYFGLTAVMCLP--KPLPEGTEDDO:ATIPSLSAML 214
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 168 GFITVYLMHNMELI-----ALIVFTLPEFLKLPK--ENNTNKAHDLDFVGLLVAT 218
QY 215 -----FLMFTPPSANSALL-----RSPIERKNAVFNTYVAVSVTAIS 254
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 219 IATVMLEFTENFNLVIMGALIAIIVFALYIKNAQRPLVNSFPQNKRYASFLEIVEM 278
QY 255 GSSLAHQGRKSKYGHSAVLP-----EGYAVG-TSCHLIPSPWLDIVGLVAGLISV 307
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 279 AIQV-----GYTFEPTEQIYHLQDPTSLLPVGYIVAVVIGALSGLT--- 324
QY 308 GAKYLPCCNVLGIPHSSINGVNSLLGLEIITYVL-VLDPVAGNGMIGFOVLL 365
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 325 -GEVLN-----SKQAITTALLSLILPAFVAGNHSIFVLSMT 364
QY 366 -----SIGELSLAIVIALTSGL-LTALLNLKIRKAP 396
Db 365 FFAGSFALMAYPLNEAIKTTIDLNMGTVAIGVYMLINVAVSGIAIAAALIDFKALNEP 424

```

```

RESULT 4
US-08-785-431-4
; Sequence 4, Application US/08785431
; Patent No. 5891667
; GENERAL INFORMATION:
; APPLICANT: Chalkef, Allison
; APPLICANT: Hodgson, John
; TITLE OF INVENTION: No. 5891667el Spoilite
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406-0939
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/785,431
; FILING DATE: 17-JAN-1997
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 9600955.0

```

```

; FILING DATE: 17-JAN-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Gimmil, Edward R
; REGISTRATION NUMBER: 38,891
; REFERENCE/DOCKET NUMBER: P31349
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-4478
; TELEFAX: 610-270-5090
; TELEX:
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 746 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-785-431-4

```

```

Query Match          4.4%; Score 93.5; DB 2; Length 746;
Best Local Similarity 24.2%; Pred. No. 0.4;
Matches 59; Conservative 37; Mismatches 87; Indels 61; Gaps 13;

```

```

QY 199 DQATIPSLSAMLGFLLMFIPSVNSALLRSPIERKNVAF-NTYVAVASV-VTAISGS 256
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 268 EEAGEVTNVSYPVPLTL-INOPAKOKATSKAEVQRGVLENTLKDFGNVAKVTOI--- 323
QY 257 SLAHQGRKSKYGHSAVLP--EGYAVGTSCHLIPSPWLDIVGLVAGLISV-GAKYLP 315
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 324 -----KIGPAVTQYELQPGVGVSKIVNLND---TALLAAK--DVRIEAPIGR 370
QY 316 CNRYLGIPIHSSIMGVNSLLGLEIITYVLVD-----TVGAGNGMIGFOVLLS 366
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 371 SAVGIEVFNENI-----SLVSLKE-----VDEKFPNSNKLVEGLGRISDPIITVP 417
QY 367 IGEISLAIVIALT-----SGLLTALLNLKIRKAPHEAKYF-----DDQVEKFP 411
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 418 LNEPHILLVAGSTGSGKVCINGITTSILLNAK---PHEVKMLIDPKKWEVLVWNGIP 473
QY 412 HLAIV 415
Db 474 HLLI 477

```

```

RESULT 5
US-08-785-431-2
; Sequence 2, Application US/08785431
; Patent No. 5891667
; GENERAL INFORMATION:
; APPLICANT: Chalkef, Allison
; APPLICANT: Hodgson, John
; TITLE OF INVENTION: No. 5891667el Spoilite
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406-0939
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/785,431
; FILING DATE: 17-JAN-1997
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 9600955.0
; FILING DATE: 17-JAN-1996
; ATTORNEY/AGENT INFORMATION:

```

NAME: Gilm, Edward R
REGISTRATION NUMBER: 38,891
REFERENCE/DOCKET NUMBER: P31349
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-270-4478
TELEFAX: 610-270-5090
TELEX:
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 788 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-785-431-2

Query Match 4.48; Score 93.5; DB 2; Length 788;
Best Local Similarity 24.28; Pred. No. 0.44;
Matches 59; Conservative 37; Mismatches 87; Indels 61; Gaps 13;

QY 199 DQATIPSLMGLFLMFRPSVNSALRSPIERKNAVF-NYYAVAVSV-VTAISGS 256
DB 310 EEAEVNVSVVPEPLL-LMOPAKOKATSKAEYQKRGQVLENTLKDFGVAKVYQI--- 365
QY 257 SLAPPOCKISKTYGSAVLP-EGYAVDTSCHLIPSPMLIVGLVAGLISYVGAKYLPQC 315
DB 366 -----KIGPVAQYVEIQPQGVKSVIYVNLHND-----IALALAAK--DVRIEAPIPGR 412
QY 316 CNRVLGIPHSSIMKYNLSLGLLEIYIYLVLD-----TVGAGNMICFOYVLS 366
DB 413 SAVGIEVPNEKI-----SLVSLKE-----VLDEKFPSSNNKLEVGIRDISGDPITVP 459
QY 367 IGEISLAIIVALT-----SGLTALLNLKIRKAPHEAKYF-----DDQVFNKFP 411
DB 460 LNEPHLVAGSTSGSKSVCINGIITSLINAK-----PHEVKMLIDPKMVELNVINGIP 515
QY 412 HLAV 415
DB 516 HLLI 519

RESULT 6
US-08-726-214-16
Sequence 16, Application US/08726214
Patent No. 6107076
GENERAL INFORMATION:
APPLICANT: Tang, Wei-Jen
APPLICANT: Gilman, Alfred G.
TITLE OF INVENTION: SOLUBLE MAMMALIAN ADENYLYL CYCLASE
TITLE OF INVENTION: AND USES THEREFOR
NUMBER OF SEQUENCES: 31
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold, White & Durkee
STREET: P.O. Box 4433
CITY: Houston
STATE: Texas
COUNTRY: United States of America
ZIP: 77210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/726,214
FILING DATE: Concurrently Herewith
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/005,498
FILING DATE: 04-OCT-1995
ATTORNEY/AGENT INFORMATION:
NAME: Highlander, Steven L.

REGISTRATION NUMBER: 37,642
REFERENCE/DOCKET NUMBER: UTSD:450
TELECOMMUNICATION INFORMATION:
TELEPHONE: (512) 418-3000
TELEFAX: (512) 474-7577
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 1248 amino acids
TYPE: amino acid
STRANDEDNESS: linear
TOPOLOGY: linear
US-08-726-214-16.

Query Match 4.38; Score 91; DB 3; Length 1248;
Best Local Similarity 23.08; Pred. No. 1.6;
Matches 45; Conservative 40; Mismatches 87; Indels 24; Gaps 6;

QY 24 LILFEFTHYDASLEDOKGLVASYQVGDITVMAAIGLFTSSFRHSMS---VAF 79
DB 191 LTLVLHLASAPMDPLKGLIGFTGIEVICALVVRKDTSHYLYGSGVTVWAM 250
QY 80 NLEPLALGVOMAILDGFLSQFSGKVITLFSIWLATMSALSVLSVDVLSKVNLAOL 139
DB 251 TQITLAAGLGYGGLGDI-----GYVLFLLFAY--SMPLPLVAILLAGLG-TSLQY 301
QY 140 VVWVLVEVTDIGNLWVINSIFNTDYMNMHIVFAAYG-----LTVMACLPKL 191
DB 302 TLQYLIRLAVFSINQVLAQV---LPMCMNTAGIFISYLSDRQORAFLETRCYEARL 358
QY 192 PEGTEDNDQATIPSL 207
DB 359 RLETFENQOEELVLSV 374

RESULT 7
US-09-031-392-6
Sequence 6, Application US/09031392
Patent No. 5942398
GENERAL INFORMATION:
APPLICANT: Tartaglia, Louis A.
APPLICANT: Weng, Xun
TITLE OF INVENTION: NUCLEIC ACID MOLECULES
TITLE OF INVENTION: ENCODING GLUTEX AND USES THEREOF
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows95
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/031,392
FILING DATE: 26-FEB-1998
ATTORNEY/AGENT INFORMATION:
NAME: Melkielejohn, Ph.D., Anita L.
REGISTRATION NUMBER: 35,283
REFERENCE/DOCKET NUMBER: 07334/072001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/542-5070
TELEFAX: 617/542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 509 amino acids
TYPE: amino acid
TOPOLOGY: linear


```

? APPLICANT: David William Holden
? TITLE OF INVENTION: Identification of Genes
? NUMBER OF SEQUENCES: 501
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: Patrea L. Pabst
? STREET: 2800 One Atlantic Center
? STREET: 1201 West Peachtree Street
? CITY: Atlanta
? STATE: Georgia
? COUNTRY: USA
? ZIP: 30309-3450
?
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: PatentIn Release #1.0, Version #1.30
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/637,759B
? FILING DATE: 03-MAY-1996
? CLASSIFICATION: 435
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: PCT/GB95/02875
? FILING DATE: 11-DEC-1995
? CLASSIFICATION: 435
? ATTORNEY/AGENT INFORMATION:
? NAME: Pabst, Patrea L.
? REGISTRATION NUMBER: 31,284
? REFERENCE/DOCKET NUMBER: RPMs 101
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (404) 873-8794
? TELEFAX: (404) 873-8795
? INFORMATION FOR SEQ ID NO: 89:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 759 amino acids
? TYPE: amino acid
? STRANDEDNESS: single
? TOPOLOGY: linear
? MOLECULE TYPE: protein
? HYPOTHEICAL: NO
? US-08-637-759B-89

```

Query Match 4.2%; Score 90; DB 2; Length 759;
 Best Local Similarity 20.5%; Pred. No. 0.97;
 Matches 75; Conservative 59; Mismatches 125; Indels 106; Gaps 17;

```

QY 52 QDLTVMAAIGLGLTSSFRHRSMS-----SVAFNLEMLALGVQNALLDGFLSQFPSC 104
   ||: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 100 QDMVLATVLLAIYVMDLPPTWVDILITINIMFSVILLIAIYLSDDPD--LSVFPSL 157
QY 105 KVVITLFSIATMSALSVLS-----VDA-----VLGKNLAQLVVMVLEVTDLGNLR 154
   ||: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 158 LITTLVRLSTLSTSKVLVLLQHNAGNIVDAFGKVVGGNLTVGLVFTITTIQF----- 213
QY 155 MVISNIENFDYHNMNMHIYFAVFGLT---VAMCLPKPLPEGTEDNDQFATIPS----- 206
   ||: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 214 IVITK-----GIEVAEVSARFSLDGMGKQMSIDGLRAGVIDADHARTLRQHVQ 265
QY 207 -----LSAMLGALFLMMPSPVNSALLSPIERKNAVFNTYAAVSVTAISGSSLAHPQ 262
   ||: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 266 ESRLFGAMDGM-----KFKVGDITAGIIVLVNIIIGGIITIAIVQ 305
QY 263 GKIS-----KTYGHSAVLPQGVAVDTSCHLIPSPWLPVIGLVAGLSVCGAKYLPG--- 314
   ||: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 306 YDMSMSAVHTY---SVLSIG---DGLCGQIPS---LLISLSAGIIVTR---VPEEKR 351
QY 315 -----CCNRYLGIPIHSSIM-----GYNFSLGLLAEIIVLVVL---DTVG 353
   ||: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 352 ONLATELSSQIARQPSQLITLAVVLMALLIPGPFITLAFFSALLAPIILIRKKSIV 411
QY 354 AGNGM 358
   ||: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 412 SANGV 416

```

```

RESULT 11
US-08-871-355A-89
? Sequence 89, Application US/08871355A
? Patent No. 6015669
? GENERAL INFORMATION:
? APPLICANT: David William Holden
? TITLE OF INVENTION: Identification of Genes
? NUMBER OF SEQUENCES: 501
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: Patrea L. Pabst
? STREET: 2800 One Atlantic Center
? STREET: 1201 West Peachtree Street
? CITY: Atlanta
? STATE: Georgia
? COUNTRY: USA
? ZIP: 30309-3450
?
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: PatentIn Release #1.0, Version #1.30
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/871,355A
? FILING DATE: 09-JUN-1997
? CLASSIFICATION: 435
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: PCT/GB95/02875
? FILING DATE: 11-DEC-1995
? CLASSIFICATION: 435
? ATTORNEY/AGENT INFORMATION:
? NAME: Pabst, Patrea L.
? REGISTRATION NUMBER: 31,284
? REFERENCE/DOCKET NUMBER: RPMs 101 CON
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (404) 873-8794
? TELEFAX: (404) 873-8795
? INFORMATION FOR SEQ ID NO: 89:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 759 amino acids
? TYPE: amino acid
? STRANDEDNESS: single
? TOPOLOGY: linear
? MOLECULE TYPE: protein
? HYPOTHEICAL: NO
? US-08-871-355A-89

```

Query Match 4.2%; Score 90; DB 3; Length 759;
 Best Local Similarity 20.5%; Pred. No. 0.97;
 Matches 75; Conservative 59; Mismatches 125; Indels 106; Gaps 17;

```

QY 52 QDLTVMAAIGLGLTSSFRHRSMS-----SVAFNLEMLALGVQNALLDGFLSQFPSC 104
   ||: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 100 QDMVLATVLLAIYVMDLPPTWVDILITINIMFSVILLIAIYLSDDPD--LSVFPSL 157
QY 105 KVVITLFSIATMSALSVLS-----VDA-----VLGKNLAQLVVMVLEVTDLGNLR 154
   ||: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 158 LITTLVRLSTLSTSKVLVLLQHNAGNIVDAFGKVVGGNLTVGLVFTITTIQF----- 213
QY 155 MVISNIENFDYHNMNMHIYFAVFGLT---VAMCLPKPLPEGTEDNDQFATIPS----- 206
   ||: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 214 IVITK-----GIEVAEVSARFSLDGMGKQMSIDGLRAGVIDADHARTLRQHVQ 265
QY 207 -----LSAMLGALFLMMPSPVNSALLSPIERKNAVFNTYAAVSVTAISGSSLAHPQ 262
   ||: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 266 ESRLFGAMDGM-----KFKVGDITAGIIVLVNIIIGGIITIAIVQ 305
QY 263 GKIS-----KTYGHSAVLPQGVAVDTSCHLIPSPWLPVIGLVAGLSVCGAKYLPG--- 314
   ||: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 306 YDMSMSAVHTY---SVLSIG---DGLCGQIPS---LLISLSAGIIVTR---VPEEKR 351

```

OY 315 -----CCNRVLGIPHSIM-----GYNFSLGLGELIYIVLVL---DTVG 353
Db 352 QNATELSQIARQPSILITLAVVLMALIPGPFITLAFPSALALPILIRKRSV 411
OY 354 AGNGM 358
Db 412 SANGV 416

RESULT 12
US-08-119-773-2

; Sequence 2, Application US/08119773
; Patent No. 5460942
; GENERAL INFORMATION:
; APPLICANT: Chou, Janice Y.
; APPLICANT: Lei, Ke-jian
; APPLICANT: Shelly, Leslie L.
; TITLE OF INVENTION: GLUCOSE-6-PHOSPHATASE: THE GENE AND
; TITLE OF INVENTION: PROTEIN AND RELATED MUTATIONS
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend Hourie and Crew
; STREET: Stewart Street Tower, One Market Plaza
; CITY: San Francisco
; STATE: California
; COUNTRY: US
; ZIP: 94105-1493
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/119,773
; FILING DATE: 10-SEP-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Weber, Kenneth A.
; REGISTRATION NUMBER: 31,677
; REFERENCE/DOCKET NUMBER: 15280-175
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 543-9600
; TELEFAX: (415) 543-5043
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 357 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-119-773-2

Query Match 4.18; Score 88; DB 1; Length 357;
Best Local Similarity 19.98; Pred. No. 0.5;

Matches 75; Conservative 52; Mismatches 116; Indels 134; Gaps 18;

OY 20 LEALILILEFFTHYDASLEDOKGLVASVOGDITVMAAIGLGLTSSFRHSSSVAF 79
Db 39 LRNAFYVLEPIWFHQA-----VGKILMWAVIG-----DMLNLVF 75
OY 80 NLEMLALGVOMAILLDGFLS-----QF-----PSGRVITLFSIWTMSAL 121
Db 76 KWILFGORPYWVLDITVYSNTSVPLIKQFPVTCETGPGSPSGHAMGTAGYVWVSTL 135
OY 122 SYL-----ISVDAVLGKVNIA-----QVYVNVL--VEYVTLG 151
Db 136 SIFGKIKPIYRERCLNVILMLGFWAVQLNVCLSRIYLAHFPPQVAVGSLAVETTF 195
OY 152 NLRVINSIENTD---YHNMNMHIYFAAYF-----GLTVANCLPK-----PLPGE 196
Db 196 S---HHSIVNASKYKFLIFLFFLFSRAIGYLLKGLGVLLWTLLEAQAQWCEQPEVNH 252
OY 197 DNDQATIP--SLSAMIGALF-LMWFPSVNSALLRSPIERKNVENVTYAVAVSVTAI 253

Db 253 ID-----TTFPASILKNIGTLELGLL---ALNMSYRSC----- 284
OY 254 SGGSLAHPOGKISK--TYGHSAVLDEGVAVGTSCILPSPWLPVILGVY-----AGLISYR 307
Db 285 -----KGKLSKWLPPRLSSIVASLVLLHVPDSLKPPSOVELVFYLSCKSAVYPLA 336
OY 308 GAKYLPGCCNRVLGIPH 324
Db 337 SVSVIPIYCLAQVLGQPH 353

RESULT 13
US-08-119-773-4

; Sequence 4, Application US/08119773
; Patent No. 5460942
; GENERAL INFORMATION:
; APPLICANT: Chou, Janice Y.
; APPLICANT: Lei, Ke-jian
; APPLICANT: Shelly, Leslie L.
; TITLE OF INVENTION: GLUCOSE-6-PHOSPHATASE: THE GENE AND
; TITLE OF INVENTION: PROTEIN AND RELATED MUTATIONS
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend Hourie and Crew
; STREET: Stewart Street Tower, One Market Plaza
; CITY: San Francisco
; STATE: California
; COUNTRY: US
; ZIP: 94105-1493
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/119,773
; FILING DATE: 10-SEP-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Weber, Kenneth A.
; REGISTRATION NUMBER: 31,677
; REFERENCE/DOCKET NUMBER: 15280-175
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 543-9600
; TELEFAX: (415) 543-5043
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 357 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 1..357
; OTHER INFORMATION: /label= R to C at 83
; OTHER INFORMATION: /note= "The G-6-Pase amino acid sequence mutated
; OTHER INFORMATION: from Arg to Cys at position 83."
; US-08-119-773-4

Query Match 4.18; Score 87; DB 1; Length 357;
Best Local Similarity 19.98; Pred. No. 0.63;

Matches 75; Conservative 52; Mismatches 116; Indels 134; Gaps 18;

OY 20 LEALILILEFFTHYDASLEDOKGLVASVOGDITVMAAIGLGLTSSFRHSSSVAF 79
Db 39 LRNAFYVLEPIWFHQA-----VGKILMWAVIG-----DMLNLVF 75


```

; INFORMATION FOR SEQ ID NO: 23
;
; SEQUENCE CHARACTERISTICS:
;     LENGTH: 603 amino acids
;     TYPE: amino acid
;     STRANDEDNESS:
;     TOPOLOGY: linear
US-09-097-889-23

```

Query Match	4.0%;	Score 85.5;	DB 4;	Length 603;
Best Local Similarity	17.4%;	Pred. No. 2;		
Matches	95;	Conservative 79;	Mismatches 177;	Indels 195; Gaps 21

```

0Y 5 YPRSVGRCLPCLATLEAAILLAEFFFNHNDASLEDDOKGVASVQVGDITLVAALIGF 64
Db 32 YPHYAK-----SIVASTIISLEPPTMCMCL-DOEVIISNMHNAFTTQTO--LSLSF 80
0Y 65 LTSSFRH-----SNSSVAFNLFMLALGOWAAILDDGFSLOPSPGKVVITLFSIWL 116
Db 81 KLDFESMSEFIPVALLFVMSIMEFSL-----WYMSDPNINFEFFYLLITLMTLIV 122
0Y 117 TMSAL-----SVLISVDVAVGLVNLALQLVNMYLVEVTGLGMLRWYISNIFN 162
Db 133 TANNLFOLFIEGEGVIMSFLILISMYARADANFAAIOAILYNRIGDIG----- 181
0Y 163 TDYHNNMMHIIYFAAYGSLVAVMCLPRPLRGESTDNDQATII--PSLSAMGLALFL-- 216
Db 182 -----FLTALAMET--LHNSNDPOOMALLNANPSLTPLIGLLLAAG 222
0Y 217 -----W--MFRPSVSNALLRS-----PIERKNAFNTYVAAVS 248
Db 223 KSAQGLGHFWLPSAMEGTTPYSALLHSTWYVAGIFILLIRHPLAEMSPLIOT-LTICLG 281
0Y 249 VVTAISGSSLAHPQOKISKTYGHSAVLpreGVAVT-----SCHLIDSPWLPVL- 297
Db 282 AITTLFAAVCALTONDIKIKIYAFSTSSOLGLMAYTIGINOPHLAFLHICHTAFKMLFM 341
0Y 298 -----GLVAGLSIVGAKLIPCCCR-----VLGI 322
Db 342 CSGSIHNLNNEODIRKMGKLLKTMPLTSTLSLTIGSLAAGMPLTGFYSKDHITETAM 401
0Y 343 PHSISMGVNSFLGLAEIIVYLVLDVT-----VGAGN 356
Db 402 SYTNMMAISITLILATSLTSAKSTKMILITLTCGQRPFLTNINNNPTLNPRIKRLAAGS 461
0Y 357 GMIGFOVLLSIG-----ELSLAIVALT-----GLTtALLN-----LKIRAPHEAK 400
Db 462 LFAGELLTNNISPAESPQTTIPLYLKLTALAVTELGILLTALDNLVNNKLKM-KSPICTF 520
0Y 401 YFDDOY 406
Db 521 YFSNML 526

```

Search completed: September 12, 2001, 07:39:08
Job time: 317 sec

THIS PAGE BLANK (USPTO)

201 GAGTTTCCGAGACACAGCTGGAGCAGTGTGACCTTCAACCTCTTCAATGC 250
 67 rasnheargatghisSerTtrpSerValAlaIaphasnLeuphemeTL 84
 251 TGGCGCTTGTCGAGTGGGCAATCCTGTGAGCGGCTTCTGAGCCAG 300
 84 euAlaLeuGlyValGIntPrAlaIleLeuLeuAspGlyPheLeuSerGln 100
 301 TTCCCTTGGGAAGTGGTGCATCACTGTTCATATTGGCTGGCCAC 350
 101 PhePrProGlyLysValValIleThrLeuPheSerIleArgLeuAlaTh 117
 351 CATGAGCTCTTGTGCTGATGTATCTGATGGATGCTGTGGGGAGG 400
 117 rMetSerIaMetSerValIleuIleSerAlaGlyAlaValLeuGlyLysV 134
 401 TCACCTTGGCGAGTGGTGTGATGTGTGTGGTGGAGGTGACACTTGA 450
 134 aIasnLeuAlaGlnLeuValValMetValLeuValGluValThrAlaLeu 150
 451 GGCAACCTGAGATGGTGCATCAGTAATCTTCAACACAGACTACACAT 500
 151 GlyThrLeuArgMetValIleSerasnIlePheasnThrAspTyrHisMe 167
 501 GAACATGATGCACATCTAGTGTTCGACGCTATTTGGCGTGTGTGG 550
 167 tasnLeuArgHisPheTyrValPheAlaIaIatyrPheGlyLeuThrValA 184
 551 CCTGTGCTGCCAAGCCTTACCAGGAGGAGGAGGAGGATTAAGATCAG 600
 184 IatPrCysLeuProLysProLeuProLysGlyThrGluAspAsnAspGln 200
 601 ACAGCAAGATACCCAGTTGTCTGCCATGCTGGCGGCTCTTCTGTG 650
 201 ArgAlaThrIleProSerLeuSerIaMetLeuGlyAlaLeuPheLeuTr 217
 651 GATGTTCTGGCCAACTTCAACTGTCTGTCTGTGAGAGTCCATCGAAA 700
 217 pMerThrProSerValAsnSerProLeuLeuArgSerProIleGlnA 234
 701 GGAGAAATGCGGTGTTCACACACTACTGTGTGACAGTACGCTGTG 750
 234 rGlyAsnAlaMetPheasnThrTyrTyrAlaLeuAlaValSerValVal 250
 751 ACAGCATCTCAGGGTGCATCTTGGCTACCCCAAGGAGATGACGCAA 800
 251 ThrAlaIleSerGlySerSerLeuAlaHisProGlnArgLysIleSerMe 267
 801 GACTTATGTGACAGTGGCTGTGGCAGAGGCGCTGGCTGTGGTACT 850
 267 tThrTyrValHisSerAlaValAlaLeuAlaGlyValAlaValAlaGlyTrs 284
 851 CGTGCACACTGATCCCTTCTCCGTGGCTGCCATGCTGCTGGCTGTG 900
 284 erCysHisLeuIleProSerProTrpLeuAlaMetValLeuGlyLeuVal 300
 901 GCTGGCTGATCTCGTGGGGGAGCCAAATACCTGCCGGGGTGTGTAA 950
 301 AlaGlyLeuIleSerIleGlyValAlaLysCysLeuProValCysCysAs 317
 951 CCGAGTGTGGGATTTCCCAAGCTCCATGAGGCTTACAACTTCACCT 1000
 317 nArgValLeuGlyIleHisHisIleSerValMetHisSerIlePheSerL 334
 1001 TGCTGGCTCTGCTGGAGAGATCATCTACATTGTGCTGCTGCTGTGAT 1050
 334 euLeuGlyLeuLeuGlyGluIleThrTyrIleValLeuLeuValLeuHis 350
 1061 ACCGTGCGAGCGCGCATGATGATGCTTCCAGGTCCTCCACAGAT 1100
 351 ThrValITrpaSngLysnGlyMetIleGlyPheGlnValLeuLeuSerI 367

1101 TGGGAACCTCAGCTTGGCCATGTCGATAGTCTGCACGCTGTGCTCTGA 1150
 367 eGlyGlnLeuSerLeuAlaIleValIleAlaLeuThrSerIleLeuLeuT 384
 1151 CAGTTTGCCTCTAAATCTTAAATATGGAAGCAACCTCATGAGGCTTAA 1200
 384 hrcGlyLeuLeuLeuAsnLeuLysIleTrpLysAlaProHisValAlaLys 400
 1201 TATTTTGTACACAAAGTTTCTGGAAGTTTCTCATTTGGCTGTGGATT 1250
 401 TyrPheAspAspGlnValPheThrLysPheProHisLeuAlaValGlyPh 417
 1251 T 1251
 417 e 417
 seq_name: /cgn2_6/prodata/2/1aa/5B_COMB.pep:US-08-483-533-41
 seq_documentation_block:
 ; Sequence 41, Application US/08483533
 ; Patent No. 6172047
 ; GENERAL INFORMATION:
 ; APPLICANT: Roizman, Bernard
 ; APPLICANT: Chou, Joany
 ; TITLE OF INVENTION: Method for Treating Tumorigenic
 ; TITLE OF INVENTION: Diseases
 ; NUMBER OF SEQUENCES: 43
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
 ; STREET: 6300 Sears Tower, 233 South Wacker Drive
 ; CITY: Chicago
 ; STATE: Illinois
 ; COUNTRY: United States of America
 ; ZIP: 60606-6402
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/483,533
 ; FILING DATE: 07-MAR-95
 ; CLASSIFICATION: 514
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 08/419,853
 ; FILING DATE: 11-APR-95
 ; PRIOR APPLICATION DATA: 07/861,233
 ; FILING DATE: 31-MAR-92
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Zeller, James P.
 ; REGISTRATION NUMBER: 28,491
 ; REFERENCE/DOCKET NUMBER: 28097/32742
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 312/474-6300
 ; TELEFAX: 312/474-0448
 ; TELEX: 25-3856
 ; INFORMATION FOR SEQ. ID NO.: 41:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 355 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-483-533-41
 alignment_scores:
 Quality: 124.00 Length: 312
 Ratio: 1.033 Gaps: 19
 Percent Similarity: 38.462 Percent Identity: 24.679
 alignment_block:

US-09-600-714-41/rev x US-08-483-533-41 ..

Align seg 1/1 to: US-08-483-533-41 from: 1 to: 355

```

921 CCCGAGGAGATCAGCCGACAGACAGCCAGCACCATTGGCAGCCACG 872
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
18 ProthrglyAlaValProthrglyAlaGlnSerGlnValThrSerThrProAs 34
    : : : : : : : : : : : : : : : : : : : : : : : :
871 GAGAGAGATCAGGTGACACGAGGTACCCACAGCAGCAGCCTCTGCAAC 822
    : : : : : : : : : : : : : : : : : : : : : : : :
34 nSerGlnProAlaVal.ArgSerAlaProAlaAlaAlaProPro.ProPr 50
    : : : : : : : : : : : : : : : : : : : : : : : :
821 ACCGCACTGTGCACATTAAGTCTTCTGTATCTTCCCTGGGGGTAGCCA. 773
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
50 oPro.....AlaGlyGlyProProp 57
    : : : : : : : : : : : : : : : : : : : : : : : :
772 .....AGATGACCCCTGAGA...TGGCTGTACACCGCTGACTGCTA 734
    : : : : : : : : : : : : : : : : : : : : : : : :
57 roSerGlySerLeuLeuLeuArgGlnThrLeu.His..... 68
    : : : : : : : : : : : : : : : : : : : : : : : :
733 CAGCATAGTAGGTGTGAACAGCGCATTCCTTCGATTGAGACTTCTC 684
    : : : : : : : : : : : : : : : : : : : : : : : :
69 .....ValPro...GluSerAlaSerA 75
    : : : : : : : : : : : : : : : : : : : : : : : :
683 AGCAGAGCAGAGTTGAACCTTGGCCAGACATCCACAAGAGAGG... 638
    : : : : : : : : : : : : : : : : : : : : : : : :
75 spAspAspAspAspAspTrpProAspSerProProGluSerAla 91
    : : : : : : : : : : : : : : : : : : : : : : : :
637 .....CG 636
    : : : : : : : : : : : : : : : : : : : : : : : :
92 ProGluAlaArgProThrAlaAlaAlaProArgProProGlyProH1sar 108
    : : : : : : : : : : : : : : : : : : : : : : : :
635 CCCAGCATGCGAGACAAACTGGGTATCGTGTGATCTTTATCTCTC 586
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
108 gProAlaTrpAlaArgGly.....AlaGlyLeuThrProp 120
    : : : : : : : : : : : : : : : : : : : : : : : :
585 CGTTCCTCTCGGGTAGAGGCTTTGGCAGGACACGCGCACAGAGCCCA 536
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
120 roThrPro.....ProArgAlaProSerAlaPhe 129
    : : : : : : : : : : : : : : : : : : : : : : : :
535 AATAGGCTGCACAGCTAGATGTCATCA..... 506
    : : : : : : : : : : : : : : : : : : : : : : : :
130 ArgArgAlaSerProSerAlaCysAlaSerProArgSerThrTrpArgAl 146
    : : : : : : : : : : : : : : : : : : : : : : : :
505 .TGTTCATGTGTAGTGTGTGAAGATATTACTGATGACATCTCGAG 457
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
146 aCysAlaCysAspAla.....ArgAlaG 154
    : : : : : : : : : : : : : : : : : : : : : : : :
456 GTTGCTTAAGCTGTACCTCCACGACGACATCCACCAACTGGCGCA 407
    || : : : : : : : : : : : : : : : : : : : : : : : :
154 LyArgGlyAlaArgSerProProArgProProArgProProArgProPro 170
    : : : : : : : : : : : : : : : : : : : : : : : :
406 AGTTGACCTTCCCAAGACAGCATCTGAGATGAGCAGCGACAAG... 359
    ||| ||| : : : : : : : : : : : : : : : : : : : : : :
171 ArgProProArgPro...ProArgGlyCysAlaSerArgProThrSerG1 186
    : : : : : : : : : : : : : : : : : : : : : : : :
358 ....CACATGATGGCGACGCGAATACTGAACAGTGTGATGACCACTT 313
    : : : : : : : : : : : : : : : : : : : : : : : :
186 yCysAlaThrTrp..... 191
    : : : : : : : : : : : : : : : : : : : : : : : :
312 CCCGAGGAGATCAGCTCAGAGAGCCGTCAGAGATTTGGCCCATGCA 263
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
192 .....SerGlyProArgProProAlaTrpArgAlaAla 202
    : : : : : : : : : : : : : : : : : : : : : : : :
262 CACCAAGGCGCAGCATGAAGAGTTGAAGGCCACACTGCTCCAGCTGTGT 213
    : : : : : : : : : : : : : : : : : : : : : : : :
203 AlaArgGlyProAla..... 207
    : : : : : : : : : : : : : : : : : : : : : : : :
212 CTCGCGAAGTCTGAGGTAGGAGGCCAAGCAATGGCCGATCAGGT 163
    ||| ||| : : : : : : : : : : : : : : : : : : : : : :
208 .SerGlyProThrGlyLeuGlySerGlyAlaGlyTrpArgArgProArgA 224
    : : : : : : : : : : : : : : : : : : : : : : : :
162 CAGATCTTGCCCACTGTAGATGACGACGAGCC..... 128
    : : : : : : : : : : : : : : : : : : : : : : : :

```

```

224 rgSerSerGlyArgAla...TrpGlyProArgProValProGlyProThr 239
    : : : : : : : : : : : : : : : : : : : : : : : :
127 CTTTGTATCCTTAAGAGAGCGTCATAGT 98
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
240 ProAlaGlnProAlaAlaArgArgThrArgSer 249
    : : : : : : : : : : : : : : : : : : : : : : : :

```

seq_name: /cgn2_6/prodata/2/1aa/PCTUS_COMB.pep:PCT-US91-06532-3

seq_documentation_block:

```

; Sequence 3, Application PC/TUS9106532
; GENERAL INFORMATION:
; APPLICANT: Roizman, Bernard
; TITLE OF INVENTION: Recombinant Herpes Simplex Viruses
; TITLE OF INVENTION: Vaccines and Methods
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
; ADDRESSEE: Bicknell
; STREET: Two First National Plaza Suite 2100
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60603
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US91/06532
; FILING DATE: 19910910
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Gruber, Lewis S.
; REGISTRATION NUMBER: 30,060
; REFERENCE/DOCKET NUMBER: 27373/8235
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/346-5750
; TELEFAX: 312/984-9740
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 355 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; PCT-US91-06532-3

```

alignment_scores:

Quality	124.00	Length:	312
Ratio:	1.033	Gaps:	19
Percent Similarity:	38.462	Percent Identity:	24.679

alignment_block:

US-09-600-714-41/rev x PCT-US91-06532-3 ..

Align seg 1/1 to: PCT-US91-06532-3 from: 1 to: 355

```

921 CCCGAGGAGATCAGCCGACAGACAGCCAGCACCATTGGCAGCCACG 872
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
18 ProthrglyAlaValProthrglyAlaGlnSerGlnValThrSerThrProAs 34
    : : : : : : : : : : : : : : : : : : : : : : : :
871 GAGAGAGATCAGGTGACACAGGTACCCACAGCAGCAGCCTCTGCAAC 822
    : : : : : : : : : : : : : : : : : : : : : : : :
34 nSerGlnProAlaVal.ArgSerAlaProAlaAlaAlaProPro.ProPr 50
    : : : : : : : : : : : : : : : : : : : : : : : :
821 ACCGCACTGTGCACATTAAGTCTTCTGTATCTTCCCTGGGGGTAGCCA. 773
    ||| ||| : : : : : : : : : : : : : : : : : : : : : :
50 oPro.....AlaGlyGlyProProp 57
    : : : : : : : : : : : : : : : : : : : : : : : :
772 .....AGATGACCCCTGAGA...TGGCTGTACACCGCTGACTGCTA 734
    : : : : : : : : : : : : : : : : : : : : : : : :

```

```

57  roSerCySerLeuLeuLeuArgInTPLeu.Hls..... 68
733 CAGCATAGTAGTGTGAAACAGCGCATCTTCCTTCATTCGACTTC 684
69  ....ValPro...GlusSerAlaSerA 75
683 AGCAGACGAGATTGAACCTGGCCAGAACATCCACAGAGAGG... 638
75  spAspAspAspAspAspSPtrProAspSerProProGluSerAla 91
637 .....CG 636
92  ProGluAlaArgProThrAlaAlaAlaProArgProProGluProAlaArg 108
635 CCCAGCATGCGAGAACAACTGGTATGCTGCTGCTGATCTTATGCTC 586
108 gProAlaTrpAlaArgGly.....AlaGlyLeuThrPro 120
585 CGTTCCTCGGCTAGAGGCTTGGCAGCAGCCAGCCAGACGCCAA 536
120 roThPro.....ProArgAlaProSerAlaPhe 129
535 AATAGGCTCGACACAGTAGATGTCATCA..... 506
130 ArgArgAlaSerProSerAlaCysAlaSerProArgSerThrTrpArgAl 146
505 .TGTTCAATGCTGTGCTGTTGAAGATATTACTGATGACCATTCCTGAG 457
146 aCysAlaCysAspAla.....ArgAlaG 154
456 GTTCGCTAAAGCTGTACCTCCACCAGCAGCATGACACCACTGCCCA 407
154 YArgGlyArgArgSerProProArgProProArgProProArgProPro 170
406 AGTTGACCTTCCCAAGACAGATCCACTGAGATGAGCAGCAGAACAG. 359
171 ArgProProArgPro...ProArgGlyCysAlaSerArgProThrSerG1 186
358 ....CACTCATGTGGGCCAGCGCAATACTGAACAAGTGTGATGACCACTT 313
186 yCysAlaThrTrpTrp..... 191
312 CCCAGAGGAGACTGGCTCAGAGAACCGTCCAGAGATTGCCCATGCA 263
192 .....SerGlyProArgProProAlaTrpArgAlaAla 202
262 CACCAAGGCGCAGCATGAAGAGTTGAAGGCCACACTGCTCAGCTGTGT 213
203 AlaArgGlyProAla..... 207
212 CTCGGAAGACTCGAGGTGAGAGAGCCCAATGCGCCCATCAGCGT 163
208 .SerGlyProThrGlyLeuGlySerGlyAlaGlyTrpArgArgProAla 224
162 CAGATCTTGGCCAACTGTATGATGCCACGAGCC..... 128
224 rgSerSerGlyArgAla...TrpGlyProArgProAlaProGlyProTrp 239
127 CCTTTTGATCTTAAGAGACCGTCATAGT 98
240 ProAlaGluProAlaArgArgTrpArgSer 249
seq_name: /cgn2_6/ptodata/2/1aa/5B_COMB.pep:US-08-635-121-2
seq_documentation_block:
; Sequence 2, Application US/08635121
; Patent No. 5910442
; GENERAL INFORMATION:
; APPLICANT: Gelman, Irwin H.
; TITLE OF INVENTION: TUMOR SUPPRESSOR GENE
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:

```

```

; ADDRESSEE: Brumbaugh, Graves, Donohue & Raymond
; STREET: 30 Rockefeller Plaza
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10112-0228
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/635,121
; FILING DATE: 19-APRIL-1996
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Clark, Richard S
; REGISTRATION NUMBER: 26,154
; REFERENCE/DOCKET NUMBER: A30558 - 165/33603
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-408-2558
; TELEFAX: 212-765-2519
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1346 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: Internal
; ORIGINAL SOURCE:
; US-08-635-121-2

alignment_scores:
Quality: 110.50 Length: 412
Ratio: 0.547 Gaps: 26
Percent Similarity: 49.029 Percent Identity: 21.602

alignment_block:
US-09-600-714-41/rev x US-08-635-121-2 ..
Align seg 1/1 to: US-08-635-121-2 from: 1 to: 1346
1155 ACCTGTCAGAGACGACGAGTGCATGATGCGATGCGCAAGCTGATT 1106
||||| : : : : :
725 ThrGlyGlnValIleLysLeuGlnGlnLeuMetGlu..GlnAlaValAl 740
1105 CCCCAGTCTGAGAG.....GACTGGAAGCCATCATCTCC.... 1069
||||| : : : : :
740 aProGluSerSerGluThrLeuThrAspSerGluThrAsnGlySerThrP 757
1068 ..ATTGCCGCTCCGACGGTATC.....AGCACCAAGCAGCAATGTA 1027
||||| : : : : :
757 roLeuAlaAspSerAspThrAlaAspGlyThrGlnGlnAspGluThrIle 773
1026 GATGATCTC...TCCAGCAGACC.....CAGCAAGCTGAAGTTGT 989
||||| : : : : :
774 AspSerGlnAspSerLysAlaThrAlaAlaValAlaValArgGlnSerGlnValTh 790
988 AGCCATGATGAGCTGTGGGAATCCCCAGCAGCTCGTTACAACACCCC 939
||||| : : : : :
790 rGluGluGlnAlaAlaThrAlaGlnGlnGluGluProSerThrLeuProA 807
938 GGCAGGACTTGCTGCCCGCCGAC...GGAGATCAGCCGACG..... 901
||||| : : : : :
807 snAsnValProAlaGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 823

```

```

900 ...CACAGACCCAGACACCAT..... 883
824 LeuGIuProThrInGIuInGIuLeuAlaAlaAlaValProValTrpGI 840
882 .....GGCAGC 876
840 nLysThrGIuValGIuInGIuInGIuValAspTrpLeuAspGIuL 857
875 CACGAGAGAGGATCAGGTGACAGAGATCCACAGCCAGCTCCGCGC 826
857 ySValLySGluInGIuInGIuVal...PheValHisSerGIuProAsnSer 872
825 CACACCGCAGCTGTGCACATAGTCTTGCTGATCTCCCTGGGGTAG 776
873 GIu...LysAlaAlaAspValThrTyraSPerGIuValMetGIuValAl 888
775 CCAAGGATGACCCCTGAGTGGCTGTACACAGCGTACGTACAGCAT. 727
888 a.....GlyCysGIuInGIuLySGluSerThrGIuValG 899
726 ....GTAGGTGTGAACAGGACATTCCTTCGATTGACCTTCAG 682
899 InSerLeuSerLeuGIuInGIuGIuMetGIuThrAsp.....ValGIu 913
681 CAGAGCAGAGTTGAACCTTGGCCAGAACATCCACAGAGAGGCGGCCA 632
914 LysGIuLySArgGIuThrLySProGIuInValSerGIuInGIuGIuGI 930
631 G.....CATGGCAGACAAACGTGGTATGCTGCTGCTC 600
930 nGIuThrAlaAlaProGIuInGIuGIuArgAsnTyGIuLySProValLeu 947
599 TGATCTTATC.....CTCCGTTCCCTGGG 574
947 hrLeuAspMetProSerSerGIuArgGIuLySAlaLeuGIuSerLeuGI 963
573 TAGAGGCTTTGGCAGGACACAGGC...CACAGACGCCCAAAATAGGCTG 527
964 .....GlySerProSerLeuProAspGIuAspLySAlaGIuCy 976
526 CGAACAGTAGATGTGCATCATGTTCATGTGTGTC..... 490
976 sILeGIuValGIuInValGIuInSerLeuAspThrValThrGIuInThrAlaG 993
489 ..TGNGTTGAAGATATT.....ACTGATGAC 466
993 LuAlaValGIuLySValILeGIuThrValAlILeSerGIuThrGIuGIu 1009
465 CATCTCAGTGTGCTTAAGCTGTACCTCCACAGCACATCACACCA 416
1010 SerProGIuLySVal...GIuAlaHisLeu..... 1018
415 ACTGGCGCAAGTTGACCTTCCCAAGACAGCATCCACTGAGATCAGCAC 366
1019 .....LeuProAlaGIu.....L 1023
365 GACAAGCAGCTCATGGTGGCCAGCGAATACTGAACAGTAGTGAGCAC 316
1023 ySSerSerAlaThrGIuLySILeGIuThrLeuGIuInHisAlaGIuAspThr 1039
315 TTTCACAGAGGAAGTGTGTGACAGACCGCTCCAGCAGATTTGCCCA. 268
1040 ValPro.....LeuGIuLySProGIuSerGIuAlaGIuSerILeProIL 1053
267 .....CTGCAC..... 262
1053 eILeValThrProAlaProGIuSerThrLeuHisProAspLeuGIuInGI 1070
261 ....ACCAAGCGCCAGCATGAGAGGTTGAAGGCCACACATGCTCCAGCT 217
1070 LuILeSerAlaSerGIuInArgGIuInArgSerGIuInGIuAspLySProAsp 1086

```

```

216 GTCTCTCGGAACTCGAGGTGAGAGCCCAAGCC 181
1087 AlaGIuProAspAlaAspLySLeuSerThrAla 1098
seq_name: /cgn2_6/prodata/2/iaa/6A_COMB.pep:US-09-121-321-16

```

```

seq_documentation_block:
; Sequence 16, Application US/09121321
; Patent No. 6090783
; GENERAL INFORMATION:
; APPLICANT: Saiga, Akihiko
; APPLICANT: Orita, Satoshi
; APPLICANT: Igataishi, Hisanaga
; APPLICANT: Okumura, Kouichi
; APPLICANT: Sakaguchi, Gaku
; TITLE OF INVENTION: DNA MOLECULE RELATING TO SUPPRESSION OF
; TITLE OF INVENTION: GENE EXPRESSION AND NOVEL PROTEIN
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FISH & NEAVE
; STREET: 1251 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10020
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/121,321
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/933,803
; FILING DATE: 19-SEP-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Haley, James F.
; REGISTRATION NUMBER: 27,794
; REFERENCE/DOCKET NUMBER: SHGN-12CIP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-596-9000
; TELEFAX: 212-596-9090
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 671 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-121-321-16

```

```

alignment_scores:
  Quality: 108.50      Length: 321
  Ratio: 0.912        Gaps: 16
  Percent Similarity: 37.072  Percent Identity: 23.053

```

```

alignment_block:
US-09-600-714-41 x US-09-121-321-16 ..

```

```

Align seg 1/1 to: US-09-121-321-16 from: 1 to: 671

```

```

42 CCTCTGGGCGCTAACACTGGAAGC.....AGCTCATTCCTCTTCT 85
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
350 ProThrAspProAsnSerGIuSerLeuLeuSerAlaHisAspLeu 366
86 ATTTT.....TACCACATATGAGCTTCCTTAGAG 117
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
366 rTripleLySGluInGIuInProTyPro..... 376
118 GATCAAAAGGGGCTGTGGCATCTATCAAGTTGC...CCAAGATCTGAC 164
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

377 .....TpglyProArgSpSer 382
165 CGTATGGCGCCATTGGCTTGGCTT..... 191
383 MetAspIyglu.....LeuGlyLeuAspSerGlyProSerAspSerLe 397
192 .....CCTCACC 198
397 uLeuMetValIySAsnProProProAlaProProGlnProGlnProGlnA 414
199 TCGAGTTCCGAGACACAGCTGGACAGTGGCTTCAACCTCTTCAT 248
414 rGlnProProGlnProGlnLeu...GlnSerGlnProGlnProGlnSer 429
249 GCTGGCCCTTGGTGGAGTGGCAATCCTGGACGGCTTCTCTGAGCC 298
430 LeuProProIleAlaValAlaGlnAsnProGlyGly.....ProProSe 444
299 AGTTCCTTCTGGAGAGTGGTTCATCACACTGTTCAGTATTTCGGTGGC 348
444 rArgGlyLeuLeuAspAspGlyPheGlnValLeuProGlyGluArgIys 461
349 ACCATGAGTGTTCGTGGTCTGCTCATCTCACT.....GGATGC 386
461 erGlyGluAlaProProGlyGlyAspArgSerThrGlyGlyGlyGly 477
387 TGCTTGGGAGAGTCAACTGGCGCATGGTGTGATGGTGGTGG 436
478 AspGlyGlyGlyGlyGlyGlyGlyAlaGlnAglThrGlyAlaGlyGly 494
437 AGGT..... 440
494 yGlyCysGlySerCysCysProGlyGlyLeuArgArgSerLeuLeuH 511
441 .....GACAGCTTGGCACTGACATGCTGATCAGTAA 476
511 IsGlyAlaArgSerLysProTyrSerCysProGlyCysGlyLysSerPhe 527
477 TATCTCAACACAGATACACATGATGATGATGATGATGATGATG 526
528 GLyAlaArgLysSerLeuIleIleHisArgSerHisThrLysGluArg 544
527 CAGCCATTGTTGGGCTGTGTGGCTGTGGCTGCAAAAGCTCTGACC 576
544 gProTyrGlyCysAlaGlyCys..... 551
577 GAGGAGACGAGAGTAAAGATCAGACAGCAACGATACCCAGTTGTCTGC 626
552 .....GluLysSerPheAsnCys 557
627 CATGCTGGGGCCCTTCTTGTGTGATGTCTGGCCAAAGTTCAACTGTG 676
558 HisSerLys.....LeuIleArgHisGlnMetH 567
677 C....TCGCTGAGAGTCCATCGAAGAGAAATGC..... 710
567 rHisArgGlyGluArgProTyrLysCysSerGlyCysGluLysThrLys 584
711 ..CGTGTCAACACCTACTATGCTGTAGCAGTCAAGCTGTGACAGCAT 758
584 erArgLysGlnHisLeu.....GlnAsnHis 592
759 CTCAGGGATCATCTTGGCTCACCCCAAGGAGATCAGCAAGACTATG 808
593 GlnArgLeuHisThrGlyGluArgProPheGln.....Cy 604
809 TGCACAGTGGCGT 821
604 sAlaLeuCysGly 608
seq_name: /cgn2_6/ptodata/2/1aa/6B_COMB.pep:US-08-933-803A-16
seq_documentation_block:

```

```

; Sequence 16, Application US/08933803A
; Patent No. 6218522
; GENERAL INFORMATION:
; APPLICANT: Saiga, Akihiko
; APPLICANT: Orita, Satoshi
; APPLICANT: Igarashi, Hisanaga
; APPLICANT: Okumura, Kouichi
; APPLICANT: Sakaguchi, Gaku
; TITLE OF INVENTION: DNA MOLECULE RELATING TO SUPPRESSION OF
; GENE EXPRESSION AND NOVEL PROTEIN
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; STREET: 1251 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10020
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/933,803A
; FILING DATE: 19-SEP-1997
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Haley, James F.
; REGISTRATION NUMBER: 27,794
; REFERENCE/DOCKET NUMBER: SHGN-12C1P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-596-9000
; TELEFAX: 212-596-9090
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 671 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-933-803A-16

alignment_scores:
      Quality: 108.50      Length: 321
      Ratio: 0.912      Gaps: 16
      Percent Similarity: 37.072      Percent Identity: 23.053

alignment_block:
US-09-600-714-41 x US-08-933-803A-16 ..
Align seg 1/1 to: US-08-933-803A-16 from: 1 to: 671
42 CCTCTGGGCCCTTAACACTGGAAGC.....AGCTTCATTCTCCTCTTCT 85
111 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
350 ProThrAspProAsnSerGlnSerLeuIleSerAlaHisAspIleLeuSe 366
86 ATTTT.....TACCACATATGACGCTTCTCTTGAAG 117
366 rTTPILeLysGlnGluGlnProTyrPro..... 376
118 GATCAAAAGGGGCTCGTGCATCTATCAAGTTGG...CCAGATCTGAC 164
377 .....TpglyProArgAspSer 382
165 CGTATGGCGCCATTGGCTTGGCTT..... 191
383 MetAspIyglu.....LeuGlyLeuAspSerGlyProSerAspSerLe 397
192 .....CCTCACC 198
397 uLeuMetValIySAsnProProProAlaProProGlnProGlnProGlnA 414

```



```

296 rSerLeuLeuGlnValThrLeuGlnValLeuLeuProArgLeuAlaValP 313
452 GCAACCTGAGATGTCATCATGATATCTTCAACACAGACTACCACATG 501
313 heserIleasnGlnValLeuAlaGlnValAla.....LeuPheMet 326
502 AACATGATGCACATCTACGTGTTCGACGCCATTGTTGGG..... 540
327 CysMetAsnThrAlaGlyIlePheIleSerTyrLeuSerAspArgAlaG 343
541 .....CTGTCTGTGGCTGTGGCTGTCCCAAGGCTCTACCCG 577
343 nArgGlnAlaPheLeuGlnThrArgArgCysValGlnAlaIaArgLeuArg 360
578 AGGACAGGAGATAAAGATCAGACAGACAGATACCAGTTGTGTGCC 627
360 eugIuThrGlnAsnGlnArgGlnIaArgLeuValLeuSerVal..... 374
628 ATGCTGGGCGCCCTCTTCTGTGGATGTTCTGGCCAAATTCAACTGTGC 677
375 .....LeuProArgPhe..... 378
678 TCTGTGAGAAATCCAAATCGAAGAGAAATGCCGTGTTCACACCTACT 727
378 ..... 378
728 ATGCTGACAGCTACGGGTGTGACAGCAATTCAGGGTCTACTGGCT 777
379 .....ValValLeuGlnMetIleAsnAspMetThrAsnValGlnAspGlu 393
778 CACCCCAAGGAGATCAGACAGCAATTATGTGCACAGT..... 816
394 HisLeuGlnHisGlnPheHisArgIleTyrIleHisArgTyrGlnAsnVal 410
817 .GCGGTGTGGCAGAGAGCGCTGTGGGTACCTCGTGCACCTGATCC 865
410 IserIleLeuPheAlaAspValIysGlyPheThrAsn..... 422
866 CTTCCTCGGCTGTGCATGGTGTGGTCTTGTGGCTGGGCTGATGTCC 915
423 .....LeuSerThr 425
916 GTCCGGGAGCCAACTACCTGCCGGGTGTGTAAACGAGTGTGGG.. 963
426 ThrLeuSerAlaGlnGlnLeuValArgMetLeuAsnGlnLeuPheAlaArg 442
964 .....ATCCCAACAGCTCATCATGGCTACAACTTCAGCTTGC 1003
442 gPheAspArgLeuAlaHisGln.....HisHisCysLeuArgIleLysIleLe 458
1004 TGGGTCTGCTGG.....AGAGATC 1023
458 uGlyAspArgTyrTyrCysValSerGlyLeuProGlnProArgGlnAspH 475
1024 ATTCATATTGTGCTGCT.....GATGCTTGA 1049
475 IAsnHisCysCysValGlnMetCylLeuSerMetIleLysThrIleAlaArg 491
1050 TACCGTGGAGCCGCGCAATGGCATGAT 1076
492 pheValArgSerArgThrLysHisAsp 500

```

seq_name: /cgn2_6/prodata/2/1aa/5b_COMB.pep:us-08-149-097D-35

seq_documentation_block:

; Sequence 35, Application US/08149097D

; Patent No. 5874236

; GENERAL INFORMATION:

; APPLICANT: Harpold, Michael

; APPLICANT: Ellis, Steven

; APPLICANT: Williams, Mark

; APPLICANT: Feldman, Daniel

; APPLICANT: McCue, Ann

```

; APPLICANT: Brenner, Robert
; TITLE OF INVENTION: HUMAN CALCIUM CHANNEL COMPOSITIONS AND
; METHOD OF INVENTION: METHODS
; NUMBER OF SEQUENCES: 40
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Brown, Martin, Haller & McClain
; STREET: 1660 Union Street
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92101-2926
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/149,097D
; FILING DATE: 05-NOV-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/105,536
; FILING DATE: 11-AUG-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/US92/06903
; FILING DATE: 14-AUG-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/914,231
; FILING DATE: 13-JUL-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/868,354
; FILING DATE: 10-APR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/745,206
; FILING DATE: 15-AUG-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/620,250
; FILING DATE: 30-NOV-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/482,384
; FILING DATE: 20-FEB-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/603,751
; FILING DATE: 04-APR-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/US89/01408
; FILING DATE: 04-APR-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/176,899
; FILING DATE: 04-APR-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: Seidman, Stephanie L.
; REGISTRATION NUMBER: 33,779
; REFERENCE/DOCKET NUMBER: 6362-55038
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 238-0062
; TELEFAX: (619) 238-0999
; INFORMATION FOR SEQ ID NO: 35:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2509 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: Internal
; ORIGINAL SOURCE:
; FEATURE:
; OTHER INFORMATION: /product="Alpha1A-1 subunit of
; human calcium channel"
; US-08-149-097D-35

```

alignment_scores:
 Quality: 106.50 Length: 382
 Ratio: 0.657 Gaps: 20
 Percent Similarity: 42.408 Percent Identity: 23.560

alignment_block:

US-09-600-714/rev x US-08-149-097D-35 ..

Align seg 1/1 to: US-08-149-097D-35 from: 1 to: 2509

```

1023 GATCTCTCCAGACAGCCAGACAGTGAAGTGTAGCC.....CANGA 980
      |||||
2185 AspleuSerMetThrThrlnSerIlyAspleuProSerIlySglAArgAs 2201
      |||||
979 TGGAGCTGTGGGAAATCCCAAG..... 958
      |||||
2201 pGlnlAArgIlyArgProIlyAspArgIlyHisArgIlnHisHisHisH 2218
      |||||
957 .....CACTCGGTTACACACCCCGGAGTACTTGCTCCGCC 919
      |||
2218 lSHlSHlSHlSHlSHlSHlSHlSHlSHlSHlSHlSHlSHlSHlSHlSHlSH 2228
      |||
918 GAGGAGATGACCCAGCCAGACAGCCAGATGCGACAGCCAGAG 869
      |||||
2229 AspIlyAspArgIlyArgIlnIlnIlnIlnIlnIlnIlnIlnIlnIlnI 2245
      |||||
868 AAGGATGAGTGACAGAGTACAGCCAGCCAGCCAGCTCCGCAACAGC 819
      |||||
2245 AArgAspGln.....ArgTTPSerArgSerProSerIlyArgIly 2259
      |||
818 GCAGTGTGCACATAGTCTGTGCTGATCTCTCCCTGGGGGTGAG..... 776
      |||
2260 .HisMetAlnHisArgIlnIlnIlnIlnIlnIlnIlnIlnIlnIlnI 2276
      |||
775 .....CCAAGATGACCCCTGAGATGCTG... 752
      |||||
2276 roSerThrSerIlyThrSerThrProArgArgIlyArgArgIlnIlnIln 2292
      |||||
751 .....TCACACGCTGACTGTCAGACAGCATAGTAGTGTGGAAC 714
      |||||
2293 GlnThrProSerThrProArgProHisValSerIlySerProValIlnI 2309
      |||||
713 ACGGATCTCTTCCTTCATTCAGATTGCACTCTCAGACAGAGCAGTTGA 664
      |||||
2309 gIysAlnIlySerIlyProProIlnIlnIlnIlnIlnIlnIlnIlnIln 2326
      |||||
663 TGGCCAGAACATCCACAAGAGAGGGCCGCGCATGGCAGCAAACTGG 614
      |||||
2326 IlnIlnIlnIlnIlnIlnIlnIlnIlnIlnIlnIlnIlnIlnIlnIlnI 2339
      |||||
613 GATCGTGTGCTGTGATCTTATCTCCGCTCCCTCGGTAGAGGCTT 564
      |||||
2340 .....GlyProArgArgIlyProIlyProIlnIlnIlnIlnIln 2352
      |||||
563 GGCAGGACAGCCAGCCAGACAGCC..... 539
      |||||
2352 uAlnIlyAspArgProProIlnIlnIlnIlnIlnIlnIlnIlnIlnIln 2369
      |||||
538 .....CAAAATAGGCTG.....CGAACAGTAGATGTGCATCA 506
      |||||
2369 rgMeGlAArgArgIlyAlProIlyProAlnArgSerIlySerProArgAl 2385
      |||||
505 TGCTATGTGTGAGTGTGTGTGAAGATATTTACTATGACCATCTCTCAG 456
      |||||
2386 CysArgHisGlyIlyAla.....ArgTTP..... 2393
      |||||
455 TTGCCTAAAGTGTACCTCCAGCAGCATACACACCACTGGCCCA 406
      |||||
2394 .....ProAlnSerIlyProHisValSerIlyG 2403
      |||||
405 GTTGACCTTCCCAAGACAGCATCTGAGATCAGACACGCAAAAGCAC 356
      |||||

```

```

2403 lypProIlyProArgHisHisIlyTyT..... 2412
355 TCATGTGGCCAGCCGAATACAGACAGTGTATACACACTTCCAGAA 306
      |||||
2413 ...TyArgIlySerAspTyT.....AspGlnAlnAs 2422
      |||||
305 GGAACCTGGCTCAGAGAGCCGTCAGACAGATTTGCCACACAGCAAG 256
      |||||
2422 pGlyProIlySerIlyGlyIlnIlnIlnIlnIlnIlnIlnIlnIlnIln 2439
      |||||
255 CCCCAGCATGAAGAGCTTGAAGCCA.....CACTGCTCCAGCTGTCTC 212
      |||||
2439 lAProPro.....ProValArgHisAlnSerSerIlyAla 2450
      |||||
211 TCCGGAACCTGAGGTGAGAGAGCCCAAGCCAAATGGCCGCAATCAGGTC 162
      |||||
2451 ThrGlyArgSer..... 2454
      |||||
161 AGATCTTGCCCACTTGATAGATGACACAGACC.....CCTTTTGATC 118
      |||||
2455 .....ProArgThrProArgAlnSerIlyProAlnIlyAla 2466
      |||||
117 CTCTAAGAGAGCTGATAGTGGGTAAAAAATAGAGAGAGATGACAG 68
      |||||
2466 lAserProSerArgHis.....GlyArgArgIlnIlnIlnIlnIln 2478
      |||||
67 CTGCTTCGAGTGTAGG.....CCACAGAGGGCAGGCGAGC 32
      |||||
2479 TyTTPProAlnHisIlyIlnIlnIlnIlnIlnIlnIlnIlnIlnIln 2493
      |||||
seq_name: /cgn2_6/plodata/2/1aa/5A_COMB.pep:us-08-313-288B-15
seq_documentation_block:
: Sequence 15, Application US/08313288B
: Patent No. 5750502
: GENERAL INFORMATION:
: APPLICANT: Jessell, Thomas M. and Avihu Klar
: TITLE OF INVENTION: CLONING, EXPRESSION AND USES OF A
: TITLE OF INVENTION: NOVEL SECRETED PROTEIN, F-SPONDIN
: NUMBER OF SEQUENCES: 20
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Cooper & Dunham LLP
: STREET: 1185 Avenue of the Americas
: CITY: New York
: STATE: New York
: COUNTRY: USA
: ZIP: 10036
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/313,288B
: FILING DATE: January 5, 1995
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: White, John P.
: REGISTRATION NUMBER: 28,678
: REFERENCE/DOCKET NUMBER: 40028-A-PCT-US
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (212) 278-0400
: TELEFAX: (212) 391-0526
: TELEX:
: INFORMATION FOR SEQ ID NO: 15:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 469 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: peptide
: HYPOTHEtical: NO
: ANTI-SENSE: NO

```



```

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/466,033
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/389,886
FILING DATE: 15-FEB-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/357,509
FILING DATE: 16-DEC-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/329,729
FILING DATE: 26-OCT-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/344,271
FILING DATE: 23-NOV-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/285,558
FILING DATE: 03-AUG-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/285,543
FILING DATE: 03-AUG-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/246,985
FILING DATE: 20-MAY-1994
ATTORNEY/AGENT INFORMATION:
NAME: Fabian, Gary R.
REGISTRATION NUMBER: 33,875
REFERENCE/DOCKET NUMBER: 4600-0201.36/G100P11
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 324-0880
TELEFAX: (415) 324-0960
INFORMATION FOR SEO ID NO: 183:
SEQUENCE CHARACTERISTICS:
LENGTH: 2910 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-466-033-183

alignment_scores:
Quality: 105.00 Length: 507
Ratio: 0.493 Gaps: 29
Percent Similarity: 42.012 Percent Identity: 21.696

alignment_block:
US-09-600-714-41 x US-08-466-033-183 ..
Align seg 1/1 to: US-08-466-033-183 from: 1 to: 2910

10 AAGTACCCGCGTGTGCGGCGCTGCGCCCTGCGGCGCTGAGCACT 59
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1440 LysAlaProAlaGlyValAlaIleGlySerIleProValIlePse..AlaVa 1455
60 GGAAGCA.....GCTCTCA 73
   ||| |||
1455 IGluAlaGlyValIleThrTrpTyrGlyMetGluProAspLeuThrAlaAsnL 1472
74 TTCCTCTTATTTTATTTTACCCACTATGACGCTTCTAGAGATCAA 123
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1472 euleuArgLeuTyrAspAspCysProTyrThrAlaAlaValAlaIleAsp 1488
124 AAGGGCGTGGCACTATCAAGTGTGCGCAAGATCTGACCGTATGAC 173
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1489 IleGlyGluAlaIleValAlaIlePheSerGly..... 1498
174 GGCCATTGGCTTGGCTTCCTACCTCGAGTTTCCGAGACAC..... 216

```

```

1499 .....LeuAlaProLeuArgMetHisProAspV 1508
217 ..ACGTGAGCAGTGTGCGCTTCAACCTTCATGCTGCGCGCTGATGTG 264
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1508 alSerTrpAlaIleValAlaArgGlyValAsnTrpProLeuLeuValGlyVal 1524
264 ..... 264
1525 GluArgThrMetCysArgGlyThrIleuSerProGlyProSerAspAspPr 1541
265 ..CAGTGGCAATCCCTGAGACGGCTTCTGACGAGTTCCTTC..... 308
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1541 oGlnTrpAla...GlyLeuIysGly..ProAsnProValProLeuLeuLe 1556
309 ....TGGAGAGTGGTCAATCACTGTCAGTATTCGCTGCGCACCAATG 354
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1556 uArgTrp.....GlyAsnAspLeuProSerIleValAlaGlyHisHis 1570
355 AGTGCTTTGTGCGTGTGATCTCACTGAGATGCTGTCTTGGGGAAGTCAA 404
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1571 .....Ile.ValAspAspLeuValArgArgGlyLeuG1 1580
405 CTGGCGCAGTTGGTGTG.....ATGTCCTGTGTG 436
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1580 yValAlaGluGlyTyrValArgCysAspAlaGlyProIleuMetValG 1597
437 AGGTGACAGCTTTAGGCAACCGTAGGATGTCATCAGTAATATGTCAC 486
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1597 IlyeAlaIleIleAlaGlyIleMet..... 1604
487 ACAGACTACACATGACATGATGATGACATCTACGTTCGACGACCTATT 536
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1605 .....IleTyrAlaSerTyrThr 1610
537 TGGCTGTCTGTGCGC.....TGG..... 555
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1610 rGlySerLeuValValIleThrAspTrpAspValIysGlyGlySerP 1627
556 .....TGGCTGCCAAGACGCTCTGACCCGAGGGA 582
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1627 rOleuTyrArgHisGlyAspGlnAlaThrProGlnProValValGln... 1642
583 ACGAGAGATAAAGATCAGACAGACAGATACCAAGTGTGTCGCATGCT 632
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1643 .....ValProProValAspHisArg.. 1649
633 GGGCGCCCTTCTGTGTGATGTTCTGGCAAGTTTCACTGCTGCTGCG 682
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1650 .....ProGlyGlyGlnSerAla.... 1655
683 TGAGAAATCCCAATCGAAAGAAATGCGCGTGTCAACACCTACTATGCT 732
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1656 .....ProSerAspAlaLys...ThrValThrAspAlaValAlaIle 1668
733 GTAGCAGTC.....AGCTGGTGAACAGCATCTGACGGTCAATC 770
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1669 IleGlnValAspCysAspTrpSerValMetThrLeuSerIleGlyGluVa 1685
771 CTGGCTCACCCCAAGGAGAGATCGAAGAACTTATGTCACAGTGGCG 819
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1685 IleuSerLeuAlaGlnAlaIleLysThrAlaGlyValAlaLysThrAla 1702
819 ..... 819
1702 yStrPLeuAlaGlyCysTyrThrGlyThrArgAlaValAlaProThrValSer 1718
820 .....GTGTGGCAGAGAGGC...GTGGCTGTGGTACCTGCTG 854
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1719 IleValAspLysLeuThrAlaGlyGlyTyrThrAlaAlaValAlaGlyHis 1735
855 TCACCTGATC.....CCTTCTC 871
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```



```

355 AGTCCTTTGTCGTCGTGATCTAGTGGAGCTCTTGGGGAAGTCAA 404
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1571 .....Ile.ValAspAspLeuValArgArgLeuGI 1580
405 CTGGCCGACAGTTGGTGNG.....ATGTCGTGGGG 436
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1580 ValAlaIaGIuGIuTyrValArgCysAspAlaGIuProIleLeuMetValG 1597
437 AGTGACAGCTTTAGGCAACCTGAGATGTCATCAGTAATATCTTCAAC 486
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1597 LylLeuAlaIleAlaGIuGIuMet.....1604
487 ACAGACTACACATGACATGATGACATCTACGTGTGCGAGCCTATT 536
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1605 .....IleTyrAlaSerTyrTh 1610
537 TGGGCTGTCTGGCC.....TGG.....555
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1610 rGIySerLeuValValAlaThrAspTTPAspValLysGIuGIuLysErp 1627
556 .....TGCCGTGCCAAGCCTCTACCGAGGGA 582
1627 rOIeUTyrArgHisGIuAspGIuAlaThrProGIuProValGIuIn... 1642
583 ACGGAGATAAAGATCAGACAGCAACGATACCAGTTTGTCTGCCATGCT 632
1643 .....ValProProValAspHisArg.. 1649
633 GGGGCCCTCTTCTTGGATGTTCTGGCCAGTTTCACTCTGCTGTCG 682
1650 .....ProGIuGIuLysErAla... 1655
683 TGAGAGTCATGGAAGAAGATGCCGTCTCAACACTACTATGCT 732
1656 .....ProSerAspAlaLys...ThrValThrAspAlaValAlaAla 1668
733 GTACCACTC.....AGCGTGTGACAGCAGCATCTAGGTCATC 770
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1669 IlegInValAspCysAspTTPSerValMetThrLeuSerIlegGIuAla 1685
771 CTGGGCTCACCCCAAGGGAAGATCAGACAGCTTATGTCACAGTGGC 819
1685 IleuSerLeuAlaGIuAlaLysThrAlaGIuAlaLysThrAlaThrAla 1702
819 .....819
1702 ySTrPLeuAlaGIuCysTyrThrGIuThrArgAlaValProThrValSer 1718
820 .....GTGTGGCAGAGAGC...GTGGCTGGGTACCTCGTG 854
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1719 IleValAspLysLeuPheAlaGIuGIuTyrPAlaAlaValAlaGIuHisC 1735
855 TCACCTGATC.....CCTTCTC 871
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1735 sHisSerValIleAlaAlaAlaValAlaAlaTyrGIuAlaSerArgSerP 1752
872 CGTGGCTTGCATGGTGGCTGTGGCTGGCTGATGTCGTCGGG 921
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1752 rOIePLeuAlaAlaAlaAlaAlaSerTyrLeuMetGIuLeu...GIuValGIu 1767
922 GCA.....GCCAAGTACCTGCCGGGTGTGTAA 950
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1768 GlyAsnAlaGIuInThrArgLeuAlaSerAlaLeuLeuGIuAlaAlaGI 1784
951 CCGAGTCTGGGGATCCCAACAGCTCATGAGGCTACAACTTC... 996
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1784 yThAlaLeuGIuThrPro.....ValValGIuLeuThrMetAlaG 1798
997 ..AGCTTCTGGTCTGCTGAGAGATCATCTACATGTCGTGGTGGTG 1044
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1798 LylAlaPheMetGIuGIuAlaSerValSerProSerLeuValThrIleu 1814

```

```

1045 CTTATACCGTCGAGCCGGCAATGCAATGATGGCTTCCAGGTCTC.. 1092
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1815 LeuGIuAlaValGIuGIuTyrPLeuGIuGIuValAlaAlaAlaSerLeuVa 1831
1093 .....CTCAGCATGGGGAACTCAGCTTGGCCA 1120
1831 lPheAspPheMetAlaGIuLysLeuSerSerGIuAspLeuTyrAlaI 1848
1121 TCGATAGCTCTCAGCTGTGCTC...CTGACAGTGTGCTCTAAAT 1167
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1848 lProValLeuThrSerProGIuAlaGIuLeuAlaGIuIleAlaLeuGI 1864
1168 CTAAATATGGAAGCA 1185
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1865 LeuValLeuTyrSerAla 1870

seq_name: /cgn2_6/prodata/2/iaa/5B_COMB.pep:US-08-464-134-183
seq_documentation_block:
: Sequence 183, Application US/08464134
: Patent No. 5849532
: GENERAL INFORMATION:
: APPLICANT: Kim, Jungsh P.
: APPLICANT: Wages, John
: APPLICANT: Young, Lavonne M.
: APPLICANT: Fry, Kirk E.
: APPLICANT: Linen, Jeffrey M.
: TITLE OF INVENTION: Hepatitis G Virus and Molecular
: TITLE OF INVENTION: Cloning Thereof
: NUMBER OF SEQUENCES: 277
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Dehlinger & Associates
: STREET: 350 Cambridge Ave., Suite 250
: CITY: Palo Alto
: STATE: CA
: COUNTRY: USA
: ZIP: 94306
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/464,134
: FILING DATE:
: CLASSIFICATION: 536
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/389,886
: FILING DATE: 15-FEB-1995
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/357,509
: FILING DATE: 16-DEC-1994
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/329,729
: FILING DATE: 26-OCT-1994
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/344,271
: FILING DATE: 23-NOV-1994
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/285,558
: FILING DATE: 03-AUG-1994
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/285,543
: FILING DATE: 03-AUG-1994
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/246,985
: FILING DATE: 20-MAY-1994
: ATTORNEY/AGENT INFORMATION:
: NAME: Fabian, Gary R.
: REGISTRATION NUMBER: 33,875
: REFERENCE/DOCKET NUMBER: 4600-0201.36/G100P11
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (415) 324-0880

```

TELEFAX: (415) 324-0960
 INFORMATION FOR SEQ ID NO: 183:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 2910 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-464-134-183

alignment_scores:
 quality: 105.00 Length: 507
 Ratio: 0.493 Gaps: 29
 Percent Similarity: 42.012 Percent Identity: 21.696

alignment block:
 US-09-600-714-41 x US-08-464-134-183 ..

Align seg 1/1 to: US-08-464-134-183 from: 1 to: 2910

```

10  AAGTACCCGGGCTGTCGCGGCGCTGCGCCCTCGGGCCCAACACT 59
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1440  LysAlaProAlaGlyValValArgSerGlyProValTyrSer...AlaVal 1455
   60  GAAGCA.....GCTCTCA 73
   |||||
1455  IGLuAlaGlyValThrTyrGlyMetGluProAspLeuThrAlaVal 1472
   74  TTCCTCTCTATTTTATACCACTATGACGCTTCCTAGAGATCAA 123
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1472  euLeuArgLeuTyrAspAspCysProTyrThrAlaAlaValAlaAsp 1488
124  AAGGCGCTCGGCGCATCATCAAGTTGGCCACATCTGACCGGATGGC 173
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1489  IleGlyGluAlaAlaValPhePheSerGly..... 1498
174  GGCCATTGGCTTGGCTTCCCTGACCTGAGTTCCGGAGACAC..... 216
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1499  .....LeuAlaProLeuArgMetHisProAspVal 1508
217  .ACGTGAGCAGTGTGGCTTCAACCTCTTCATGCTGGCGCTTGGTGTG 264
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1508  alSerThrAlaValArgGlyValAlaSerProLeuLeuValGlyVal 1524
264  ..... 264
1525  GlnArgThrMetCysArgGluThrLeuSerProGlyProSerAspAspPr 1541
265  .CACTGGGCAATCTCTGCGACGGCTTCTGAGCCAGTTCCTTC..... 308
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1541  ogLTrpAla...GlyLeuLysGly..ProAsnProValProLeuLeuLe 1556
309  ....TGGGAAGGTGTCATCACACTGTTCACTATTGCGCTGGCCACCATG 354
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1556  uArgTyrP.....GlyAsnAspLeuProSerLysValAlaGlyHisHis 1570
355  AGTCCTTTGCGGTCGATCTCAGTGTGATGCTGCTTGGGAGAGGCAA 404
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1571  .....Ile..ValAspAspLeuValArgArgLeuGlu 1580
405  CTTCGGCCAGTTGGTGGTGC.....ATGGTGGTGGTG 436
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1580  yValAlaGluGlyTyrValArgCysAspAlaGlyProIleLeuMetValG 1597
437  AGGTGACAGCTTAGGCAACCTGAGATGGATGCATCACTAAATATCTTCAAC 486
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1597  LysLeuAlaIleAlaGlyGlyMet..... 1604
487  ACAGACTACCAATGAAATGATGACATCTAGCTGTTCGAGGCTATTT 536
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1605  .....IleTyrAlaSerTyrThr 1610
537  TGGGCTGTCTGTGGCC.....TGG..... 555
```

```

1610  rgLysSerLeuValValThrAspTrpAspValLysGlyGlySerP 1627
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
556  .....TGCCTGCCAAGCCTTCACCCGAGGA 582
1627  roLeuTyrArgHisGlyAspGlnAlaThrProGlnProValGln... 1642
583  ACGAGAGATAAGATCAGACAGACATCCAGTGTGCTGCCATGCT 632
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1643  .....ValProValAspHisArg.. 1649
633  GGGGCGCCTCTCTTGTGATGTTTGGCCAGTTTCACTGCTCTGTC 662
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1650  .....ProGlyGlyGluSerAla.... 1655
683  TGAGACAGTCCATCGAAGAGAAATGCCGTTCACACACTACTATGCT 732
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1656  .....ProSerAspAlaLys...ThrValThrAspAlaValAlaAla 1668
733  GTAGCAGTC.....ACGTGTGACAGCCATCTCAGGTCATC 770
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1669  IleGlnValAspCysAspTrpSerValMetThrLeuSerIleGlyGluVal 1685
771  CTTCGCTCACCCCAAGGAGATCAGCAAGACTTATGTCACAGTGGC 819
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1685  IleuSerLeuAlaGlnAlaLysThrAlaGluAlaTyrThrAlaThrAla 1702
819  ..... 819
1702  yStrLeuAlaGlyCysTyrThrGlyThrArgAlaValProThrValSer 1718
820  .....GTGTGGCAGAGGC...GTGCTGTGGTACCTCTGTG 854
1719  IleValAspLysLeuPheAlaGlyGlyTyrAlaAlaValAlaGlyHisC 1735
855  TCACCTGATC.....CCTTCTC 871
1735  shLysSerValIleAlaAlaAlaValAlaAlaTyrGlyAlaSerArgSerP 1752
872  CGTGGCTTGGCATGCTGCTGGGCTTGTGGCTGAGTATCCGCTGGG 921
1752  roProLeuAlaAlaAlaAlaAlaSerTyrLeuMetGlyLeu...GlyValGly 1767
922  GGA.....GCCAAGTACCTCGCGGCTGTGTAA 950
1768  GlyAsnAlaGlnThrArgLeuAlaSerAlaLeuLeuLeuGlyAlaAlaG 1784
951  CCGAGTGTGGGATTCGCCACAGCTTCATCAGGCTACACTTC..... 996
1784  yThrAlaLeuGlyThrPro.....ValValGlyLeuThrMetAlaG 1798
997  .ACGTGCTGGGTCTGCTTGGAGAGATCATCTACATTTGCTGCTGGTG 1044
1798  LysAlaPheMetGlyGlyAlaSerValSerProSerLeuValThrIleLeu 1814
1045  CTTGATACCTCGAGCGCGCAATGGCATGATTGGCTTCACAGTCTC.. 1092
1815  LeuGlyAlaValAlaGlyGlyTyrProGluGlyValValAsnAlaAlaSerLeuVal 1831
1093  .....CTCAGCAATTGGGGAACCTGAGCTTGGCCA 1120
1831  LpheaSppPheMetAlaGlyLysLeuSerSerGluAspLeuTyrThrAla 1848
1121  TCGTATAGCTCTCAGCTGTGCTC...CTGACAGGTTGTCTCTTAAT 1167
1848  LeproValIleuThrSerProGlyAlaGlyLeuAlaGlyIleAlaLeuGly 1864
1168  CTTAAATATGAAAGCA 1185
1865  LeuValLeuTyrSerAla 1870
```

seq_name: /cgn2_6/ptodata/2/1aa/5B_COMB.pep:US-08-461-361-183


```

seq_documentation_block:
: Sequence 183, Application US/08461361
: Patent No. 5856134
: GENERAL INFORMATION:
: APPLICANT: Kim, Jungsh P.
: APPLICANT: Wages, John
: APPLICANT: Young, Lavonne M.
: APPLICANT: Fry, Kirk E.
: APPLICANT: Linmen, Jeffrey M.
: TITLE OF INVENTION: Hepatitis G Virus and Molecular
: TITLE OF INVENTION: Cloning Thereof
: NUMBER OF SEQUENCES: 277
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Dehlinger & Associates
: STREET: 350 Cambridge Ave., Suite 250
: CITY: Palo Alto
: STATE: CA
: COUNTRY: USA
: ZIP: 94306
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/461,361
: FILING DATE:
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/389,886
: FILING DATE: 15-FEB-1995
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/357,509
: FILING DATE: 16-DEC-1994
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/329,729
: FILING DATE: 26-OCT-1994
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/344,271
: FILING DATE: 23-NOV-1994
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/285,558
: FILING DATE: 03-AUG-1994
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/285,543
: FILING DATE: 03-AUG-1994
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/246,985
: FILING DATE: 20-MAY-1994
: ATTORNEY/AGENT INFORMATION:
: NAME: Fabian, Gary R.
: REGISTRATION NUMBER: 33,875
: REFERENCE/DOCKET NUMBER: 4600-0201.36/G100P11
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (415) 324-0880
: TELEFAX: (415) 324-0960
: INFORMATION FOR SEQ ID NO: 183:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 2910 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: US-08-461-361-183

```

```

alignment_scores:
Quality: 105.00 Length: 507
Ratio: 0.493 Gaps: 29
Percent Similarity: 42.012 Percent Identity: 21.696

```

```

alignment_block:
US-09-600-714-41 x US-08-461-361-183

```

```

Align seg 1/1 to: US-08-461-361-183 from: 1 to: 2910
10 AAGTACCCGCGGTCTGTCGGCGCTGCCCTGCGCCCTGAGCCCTTAACACT 59
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1440 LysAlaProAlaGlyValValArgSerGlyProValTrrPser...AlaVal 1455
60 GGAGCA.....GCTGCA 73
1455 lGlnAlaGlyValThrTrpTyrGlyMetGlnProAspLeuThrAlaAsnL 1472
74 TTCCTCTCTCTATTTTTCACCATATGACGCTTCCTCTAGAGATCAA 123
: ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
1472 euLeuArgLeuTyrAspAspCysProTyrThrAlaAlaValAlaAla 1488
124 AAGGCGCTGTCGATCTTCATCAAGTGGCCCAAGATCTGACCGTATGCG 173
||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
1489 lIleGlyGlnAlaAlaValAlaPhePheSerGly..... 1498
174 GGCATTTGGCTTGGCTTCCTCACCCTCGAGTTTCCGAGACAC..... 216
1499 .....LeuAlaProLeuArgMetHisProAspV 1508
217 ..AGCTGAGCAGTGTGGCTTCAACCTCTCATGCTGCGCTGTGTG 264
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1508 alSerTrpAlaLysValArgGlyValAsnTrpProLeuLeuValGlyVal 1524
264 ..... 264
1525 GlnArgThrMetCysArgGlnThrLeuSerProGlyProSerAspAspPr 1541
265 ..CAGTGGCAATCTGCTGGACGCGCTTCCTGAGCCAGTTCCCTTC..... 308
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1541 cGlnTrpAla...GlyLeuLysGly..ProAsnProValProLeuLeuLe 1556
309 ....TGGGAGGTGTCATCAACACGTTCACATTCGCTGGCCACCATG 354
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1556 uArgTrp.....GlyAsnAspLeuProSerLysValAlaGlyHisHis 1570
355 AGTCGTTGTGCGTGCATCTCATGATGATGCTGCTTGGGAGAGTCAA 404
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1571 .....Ile.ValAspAspLeuValArgArgLeuGl 1580
405 CTGGCGCAGTGTGTG.....ATGTCGTGTGTG 436
: ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
1580 yAlaAlaGlnGlyTyrValArgCysAspAlaGlyProIleLeuMetValG 1597
437 AGGTACAGCTTTAGCAACCTGAGGATGTCATCAGTAATATCTTCAAC 486
: ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
1597 lLeuAlaAlaLeuAlaGlyMet..... 1604
487 ACAGACTACACATGACATGATGACATCTACGTCGTTCGACGCTATT 536
: ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
1605 .....IleTyrAlaSerTyrThr 1610
537 TGGGCTGTCTGTGGCC.....TGC..... 555
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1610 rGlySerLeuValValValThrAspTrpAspValLysGlyGlySerP 1627
556 .....TGCTGCCAAGCCTTACCAGGGA 582
: ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
1627 rGlySerLeuValValValThrAspTrpAspValLysGlyGlySerP 1642
583 ACGGAGATTAAGATCAGACAGCAACGATACCCAGTTGTGTCATGCT 632
: ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
1643 .....ValProValAlaSerHisArg... 1649
633 GGGCGCCCTCTTCTTGATGATCTTGGCCAGTTTCAACTGCTGTGC 682
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1650 .....ProGlyGlyGlnSerAla.... 1655
683 TGAGAAGTTCATCGAAGAGAGATGCGCTGTTCACACCTACTATGCT 732
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1656 .....ProSerAspAlaLys...ThrValThrAspAlaValAlaAla 1668

```



```

174 GGCATTGGCTGGGCTTCCTCACTGAGTTCCGAGACAC..... 216
      ::::: |||
1499 .....LeuAlaProLeuArgMethisProAspV 1508
217 ..ACGTGAGCAGTGTGGCTTCAACCTCTTCAGTCGGCGCTGGTGTG 264
      ||||| ::::: ||| ::::: |||
1508 aISerTPAlaLysValArgLysValAsnTPProLeuLeuValGlyVal 1524
264 ..... 264
1525 GlnArgThiMetCysArgGluThiLeuSerProGlyProSerAspAspPr 1541
265 ..CAGTGGGCAATCCTGTGAGCGGCTTCAGCCAGTTCCTCC..... 308
      ||||| ||| ||| ::::: |||||
1541 oGlnTPAla...GlyLeuLysGly..ProAsnProValProLeuLeuLe 1556
309 ....TGGGAAGTGTGATCAGTCTGATTCAGTATTCGGCTGGCCACCTG 354
      ||| ||||| ::::: |||||
1556 uArgTPR.....GlyAsnAspLeuProSerLysValAlaGlyHisHis. 1570
355 AGTCTTGTGCTGCTGATCTCAGTGTGCTGCTGGGAGAGTCAA 404
      :: ||||| ::::: :::::
1571 .....Ile.ValAspAspLeuValArgArgLeuG1 1580
405 CTGAGCGCAGTTGGTGTG.....ATGCTGTGTGTG 436
      ::::: |||
1580 yValAlaGluGlyValArgCysAspAlaGlyProIleLeuMetValG 1597
437 AGGTGACAGCTTGAAGCAACCTGAGAGTGCATCAGTAATATCTTCAC 486
      ::::: |||
1597 LyLeuAlaIleAlaGlyLysMet..... 1604
487 ACAGACTACACATGACATGATGACATCTACGTGTGCGAGCCTATT 536
      ::::: |||||
1605 .....IleTyrAlaSerTyrTh 1610
537 TGGCTGTGTGTGCC.....TGG..... 555
      ||| |||||
1610 rGlySerLeuValValAlaThiAspTyrAspValLysGlyLysSerP 1627
556 .....TGCCGCCAAGCCTTCACCCGAGGA 582
      ||||| ::::: |||||
1627 rOleuTyrArgHisGlyAspGlnAlaThiProGlnProValValGln... 1642
583 ACGAGATTAAGATACAGACAGCAACGATACAGTTGTGTGCCATGCT 632
      ::::: |||
1643 .....ValProProValAspHisArg.. 1649
633 GGGGCGCCCTTCTTGATGTTCTGGCAAGTTTCACTGTGCTGTGC 682
      ||||| ::::: |||||
1650 .....ProGlyGlyLysSerAla..... 1655
683 TGAGACATCCATGAAGAAGATGCCGTTCACACCTACTATGCT 732
      ||| ::::: ||| ::::: |||
1656 .....ProSerAspAlaLys...ThiValThiAspAlaValAlaAla 1668
733 GTACAGATC.....ACGTGTGACAGCATCTCAGGCTCATC 770
      :: ||| ||||| ::::: |||
1669 IleGlnValAspCysAspTPSerValMetThiLeuSerIleGlyGluVa 1685
771 CTGTGGCTCACCCCAAGGAAGATCAGACAGACTTATGTGACAGTGGC 819
      ||||| ||||| ::::: |||||
1685 lleuSerLeuAlaGlnAlaLysThiAlaGluAlaTyrThiAlaThiAla 1702
819 ..... 819
1702 ySTrPLeuAlaGlyCysTyrThiGlyThiArgAlaValProThiValSer 1718
820 .....GTGTGGCAGAGAGC...GTGGCTGTGGTACCTCGG 854
      ::::: |||||
1719 IleValAspLysLeuPheAlaGlyLysTyrPAlaAlaValAlaGlyHis 1735
855 TCACCTGATC.....CCTTCTC 871

```

```

1735 shISerValIleAlaAlaValAlaAlaLysGlyAlaSerArgSerP 1752
      ||| ::::: |||
872 CGTGCGTTCATGCTGGCTGTGGCTGTGGCTGAGCTATCCGTGAGG 921
      || ||||| ::::: ||| |||||
1752 rOProLeuAlaAlaAlaAlaSerTyrLeuMetGlyLeu...GlyValG1 1767
922 GGA.....GCCAAGTACCTGCCGGGCTGTGTA 950
      ||| ::::: ||| |||||
1768 GlyAsnAlaGlnThiArgLeuAlaSerAlaLeuLeuGlyAlaAlaG1 1784
951 CCGAGTGTGGGATTCGCCACAGCTCCATCATGAGGCTACACTTC... 996
      :: ||||| ||| ::::: |||
1784 yThiAlaLeuGlyThiPro.....ValValGlyLeuThiMetAlaG 1798
997 ..ACGTGCTGGTCTGCTGTGAGAGATCATCTACATGTGCTGTGTG 1044
      ::::: |||||
1798 LyAlaPheMetGlyGlyAlaSerValSerProSerLeuValThiLeu 1814
1045 CTGTATACCTGCGAGCCGCAATGGCATGTGCTTCACAGTCTTC.. 1092
      ||| ::::: |||||
1815 LeuGlyAlaValGlyGlyTyrGlyGlyValValAlaAlaAlaSerLeuVa 1831
1093 .....CTCAGCATTTGGGCACTCAGCTTGGCCA 1120
      ||||| |||||
1831 lPheAspPheMetAlaGlyLysLeuSerSerGlyAspLeuTyrPyrAlaI 1848
1121 TCGTATAGCTCTCAGCTGTGCTC...CTGACAGTGTGCTCTANAT 1167
      ||| ::::: |||
1848 lProValLeuThiSerProGlyAlaGlyLeuAlaGlyIleAlaLeuGly 1864
1168 CTAAATATGAAAGCA 1185
      || ::::: |||
1865 LeuValLeuTyrSerAla 1870

seq_name: /cgn2_6/prodata/2/iaa/PCTUS_COMB.pep:PCT-US95-06266-157
seq_documentation_block:
; Sequence 157, Application PC/TUS9506266
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: Detection of Viral Antigens Coded
; TITLE OF INVENTION: by Reverse Reading Frames
; NUMBER OF SEQUENCES: 157
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: 350 Cambridge Avenue, Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/06266
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/246,985
; FILING DATE: 20-MAY-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/285,561
; FILING DATE: 03-AUG-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/329,729
; FILING DATE: 26-OCT-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/344,271
; FILING DATE: 23-NOV-1994
; PRIOR APPLICATION DATA:

```



```
1121 TCGTGAATAGCTCTCAGCTGTGGTCTC...CTGACAGTTTGCTCCTAAAT 1167
||::: ||| ||::: |||
1848 IeProValLeuThrSerProGlyAlaGlyLeuAlaGlyIleAlaLeuGly 1864
1168 CTTAAATATATGGAAGCA 1185
||::: |||
1865 LeuValLeuTyrSerAla 1870
```

THIS PAGE BLANK (USPTO)

OM of: US-09-600-714-41 to: SPTREMBL_16.* out_format : pfs
Date: Sep 11, 2001 8:52 AM
About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:

-MODEL=frame+ncp.model -DEV=xlp
-O=/cgn2.1/USPTO.spool/US09600714/runat.11092001.083824.22159/app_query.fasta.1.1332
-DB=SPTREMBL_16 -QEMT=fastan -SUFFIX=rspt -GAPOP=12.000
-GAPEXT=4.000 -MINATCH=0.100 -LOOPEL=0.000 -LOOPEXT=0.000
-GAPOP=4.500 -GAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500
-GAPOP=6.000 -GAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500
-DELCP=6.000 -DELEXT=7.000 -START=1 -MATRIX=blomsu62
-TRANS=human40.cdi -LIST=45 -DOCLIGN=200 -THR_SCORE=pct
-THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFM=pfis
-NORM=ext -MINLEN=0 -MAXLEN=200000000
-USER=US09600714@cgn1.1.133 -NCPU=6 -ICPU=3 -LONGLOG -NO_XLPUX
-WAIT -THREADS=1

Search information block:
Query: US-09-600-714-41

Query length: 1254
Database: SPTREMBL_16.*
Database sequences: 425026
Database length: 13205027
Search time (sec): 69.770000

score_list:

Sequence	Strd Orig	Zscore	Escore	Len	Documentation
sp_human:09U020	+	2119.00	2531.60	2.1e-133	09u020 homo sapiens (human). rf
sp_human:09U022	+	2117.00	2529.21	2.9e-133	09u022 homo sapiens (human). rf
sp_human:09U023	+	2115.00	2526.82	3.9e-133	09u023 homo sapiens (human). rf
sp_human:09U021	+	2113.00	2524.43	5.3e-133	09u021 homo sapiens (human). rf
sp_human:09U024	+	2107.00	2517.27	1.3e-132	09u024 homo sapiens (human). rf
sp_human:09U025	+	2104.00	2513.68	2.1e-132	09u025 homo sapiens (human). rf
sp_human:09U026	+	2094.00	2501.74	9.7e-132	09u026 homo sapiens (human). rf
sp_human:09U027	+	2094.00	2501.74	9.7e-132	09u027 homo sapiens (human). rf
sp_human:09U028	+	2043.00	2440.82	2.4e-128	09u028 homo sapiens (human). rf
sp_human:09U029	+	2039.00	2436.04	4.4e-128	09u029 homo sapiens (human). rf
sp_human:09U030	+	2033.00	2428.88	1.1e-127	09u030 homo sapiens (human). rf
sp_human:09U031	+	2016.00	2408.57	1.5e-126	09u031 homo sapiens (human). rf
sp_human:09U032	+	1963.00	2345.26	5.1e-123	09u032 homo sapiens (human). rf
sp_human:09U033	+	1940.00	2317.79	1.7e-121	09u033 homo sapiens (human). rf
sp_human:09U034	+	1928.00	2303.46	1.1e-120	09u034 homo sapiens (human). rf
sp_human:09U035	+	1888.00	2255.98	5.0e-118	09u035 homo sapiens (human). rf
sp_human:09U036	+	1860.00	2232.63	3.3e-116	09u036 homo sapiens (human). rf
sp_human:09U037	+	1652.00	1973.94	2.5e-102	09u037 homo sapiens (human). rf
sp_human:09U038	+	1566.00	1871.22	1.3e-96	09u038 homo sapiens (human). rf
sp_human:09U039	+	1545.00	1846.16	3.3e-95	09u039 homo sapiens (human). rf
sp_human:09U040	+	1541.00	1841.39	6.1e-95	09u040 homo sapiens (human). rf
sp_human:09U041	+	1537.00	1836.66	1.1e-94	09u041 homo sapiens (human). rf
sp_human:09U042	+	1536.00	1835.41	1.3e-94	09u042 homo sapiens (human). rf
sp_human:09U043	+	1509.00	1803.16	8.2e-93	09u043 homo sapiens (human). rf
sp_human:09U044	+	1321.00	1579.60	2.7e-80	09u044 homo sapiens (human). rf
sp_human:09U045	+	1235.00	1475.69	1.4e-74	09u045 homo sapiens (human). rf
sp_human:09U046	+	1149.00	1373.55	6.7e-69	09u046 homo sapiens (human). rf
sp_human:09U047	+	1144.00	1367.58	1.4e-68	09u047 homo sapiens (human). rf
sp_human:09U048	+	1142.00	1365.19	2.0e-68	09u048 homo sapiens (human). rf
sp_human:09U049	+	1131.00	1352.05	1.1e-67	09u049 homo sapiens (human). rf
sp_human:09U050	+	1105.00	1333.83	6.8e-66	09u050 homo sapiens (human). rf
sp_human:09U051	+	1089.00	1302.08	6.6e-65	09u051 homo sapiens (human). rf
sp_human:09U052	+	1087.00	1299.43	9.0e-65	09u052 homo sapiens (human). rf
sp_human:09U053	+	964.00	1155.13	1.6e-56	09u053 homo sapiens (human). rf
sp_human:09U054	+	801.00	963.26	1.2e-45	09u054 homo sapiens (human). rf
sp_human:09U055	+	733.00	878.33	3.8e-41	09u055 homo sapiens (human). rf
sp_human:09U056	+	730.00	879.76	6.4e-41	09u056 homo sapiens (human). rf
sp_human:09U057	+	646.00	779.95	2.6e-35	09u057 homo sapiens (human). rf
sp_human:09U058	+	638.00	753.32	6.7e-35	09u058 homo sapiens (human). rf
sp_human:09U059	+	633.00	753.32	1.4e-34	09u059 homo sapiens (human). rf

sp_mammal:09G062	+	632.50	755.86	1.7e-34	428	09g062 macaca mulatta (rhesu
sp_mammal:077484	+	631.50	754.66	2.0e-34	428	077484 macaca fascicularis (
sp_mammal:09G061	+	629.50	753.47	2.3e-34	428	09g061 papio hamadryas (hama
sp_mammal:09G065	+	629.50	752.57	2.7e-34	409	09g065 gorilla gorilla (gori
sp_mammal:09G064	+	628.50	751.08	3.1e-34	428	09g064 pongo pygmaeus (orang

seq_name: sp_human:09U020

seq_documentation_block:

ID 09U020 PRELIMINARY; PRT; 417 AA.
AC 09U020;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)
DE RH BLOOD GROUP D ANTIGEN (RHD).
GN RHDYO.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=YO;
RA Hyodo H., Ishikawa Y., Kashiwase K., Ogawa A., Watanabe Y.,
RA Tsuneyama H., Toyoda C., Uchikawa M., Akaza T., Fujii T.;
RT "Polymorphisms of RHDya in Japanese."
RT Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB018969; BA82159.1; -
DR InterPro; IPR001905; -
DR Pfam; PF00909; Ammonium_transp; 1.
DR PRINTS; PR00342; RHESUSRHD.
SQ SEQUENCE 417 AA; 45179 MW; 62B11DD24064F7D5 CRC64;

alignment_scores:
Quality: 2119.00 Length: 417
Ratio: 5.082 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 99.760

alignment_block:
US-09-600-714-41 x 09U020 ..

Align seg 1/1 to: 09U020 from: 1 to: 417

1	AGAGCTCTTAAGTACCGGCTGCTGCGGCGCTGCGCCCTGAGGC	50
1	MetSerSerLysTrpProArgSerValArgArgCysLeuProLeuTrpAl	17
51	CCTAACACTGGAAGCAGCTCTCATCTCTCTTCATTTTATACCACT	100
17	AlauThrLeuGluAlaAlaLeuLeuLeuLeuPheTrpPheTrpHisT	34
101	ATGACGCTTCCTAGAGATCAAAAGGCGCTGCGCATCTATCAAGTT	150
34	yAlaSerAlaSerLeuGluGlnSpGlnLysGlyLeuValAlaSerTyrGlnVal	50
151	GGCCCAAGATCAGCGGTATGCGGCGCATTTGGCTTGGCTTCACAC	200
51	GAGTTCGCGAGACACAGCTGAGAGAGTGGCTTCAACCTTCATGCG	250
67	rSerPheArgArgHisSerTrpSerSerValAlaPheAsnLeuPheMetL	84
251	TGGCGCTTGGTGTCAGAGTGGCAATCTGCTGGAGCGCTTCTGAGCCAG	300
84	eValAlaLeuLysValGlnTrpAlaLeuLeuLeuAspLysPheLeuSerGln	100
301	TTCGCTTCGGAGAGGTGTCATCACAGCTTCAGATTCGGCTGCGCCAG	350
101	PheProSerGlyLysValValIleThrLeuPheSerIleArgLeuAlaTrh	117

351 CATGAGTCCTTTGCGTGCATCTCAGTGCAGTGCCTGTTGGGGAAG 400
 117 rmetseralaleuSERValleuIleSerValAspAlaValleuGlyLysV 134
 401 TCACTTGGCCAGTGGTGGTGAAGTGGTGGTGGAGGTACAGCTTAA 450
 134 aAsnLeuAlaGlnValValMetValLeuValGlnValThrAlaLeu 150
 451 GGCAACCTGAGAGATGCTATCATGTAATATCTTCAACACAGACTACCAT 500
 151 GlysAlaLeuAlaGlyMetValIleSerAlaIlePheAsnThrAspTyrHisMe 167
 501 GAACATGATGACATCTACGTGTTCGACGCTATTTGGCTGTCTGG 550
 167 tAsnMetHisIleTyrValAlaPheAlaIleTyrPheGlyLeuSerVal 184
 551 CCTGGTGGCTGCCAAGGCTCTACCCGAGGAGACGAGAGATTAAGATAG 600
 184 latrPCysLeuProLysProLeuProGluGlyThrGluAspLysAspGln 200
 601 ACAGCAAGATACCCAGTTTGTCTCCATGCTGGGCGCCCTTTCTTGTG 650
 201 ThrAlaThrIleProSerLeuSerAlaMetLeuGlyAlaLeuPheLeuTr 217
 651 GATGTTCTGGCCAAAGTTTCAACTGTCTGTCTGTGAAGATCCAAATCGAA 700
 217 pMetPheTrpProSerPheAsnSerAlaLeuLeuAspSerProIleLysA 234
 701 GGAGAAATGCCGCTGTCAACACCTTCTATGCTGTAGCAGTACGCTGTG 750
 234 rGlyAsnAlaValAlaPheAsnThrTyrTyrAlaValAlaValSerVal 250
 751 ACAGCCATCTCAGAGGTCAATCTTGGCTCACCACCCCAAGGAGATCAGCA 800
 251 ThrAlaIleSerGlySerSerLeuAlaHisProGlnGlyLysIleSerLys 267
 801 GACTTATGTGCACAGTGGGTGGTGGCAGAGGCGCTGGCTGGTACT 850
 267 sThrTyrAlaHisSerAlaValAlaLeuAlaGlyGlyAlaValAlaGlyThrs 284
 851 CGTGCACCTGATCCCTTCCGTTGGCTTGGCATGGTCTGGCTTGTG 900
 284 erCysHisLeuIleProSerProTrpLeuAlaMetValLeuGlyLeuVal 300
 901 GCTGGGCTGATCTCCGTCGGGGAGCCAAATACCTCGGGGCTGTGTA 950
 301 AlaGlyLeuIleSerValGlyGlyAlaLysTyrLeuProGlyCysCysAs 317
 951 CCGAGTCTGGGATTCSSCACAGCTCCATCATGGGCTACAACTCAGCT 1000
 317 nArgValLeuGlyIleProHisSerSerIleMetGlyTyrAsnPheSerL 334
 1001 TGTGGGTCTGCTTGGAGAGATCATCTACATTTGCTGCTGCTTAT 1050
 334 euLeuGlyLeuLeuGlyGlnIleIleTyrIleValLeuLeuValLeuAsp 350
 1051 ACCGTCGAGCGGCAATGCGCATGTTGGCTTCCAGTCTCCATCAGAT 1100
 351 ThrValGlyAlaGlyAsnGlyMetIleGlyPheGlnValLeuLeuSerIle 367
 1101 TGGGAACTCAGCTTGGCCATCGTATAGCTCTCACTGCTGCTCTGA 1150
 367 eGlyIleuSerLeuAlaIleValIleAlaLeuThrSerGlyLeuLeu 384
 1151 CAGGTTCTCTAAATCTTAAATATGAAAGACCTCATGAGGCTAAA 1200
 384 hrGlyLeuLeuLeuAsnLysIleTyrAlaAlaProHisGlnAlaLys 400
 1201 TATTTTGAACAAGTTTCTGGAAGTTTCTGATTTGGCTGGTGAAT 1250
 401 TyrPheAspAspGlnValAlaPheTrpLysPheProHisLeuAlaValGlyPh 417
 1251 T 1251

seq_name: sp_human:09UQ22

417 e 417

seq_documentation_block:
 ID 09UQ22 PRELIMINARY; PRT; 417 AA.
 AC 09UQ22;
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)
 DE RH BLOOD GROUP D ANTIGEN (RHD).
 GN RHDVA.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OC NCBL_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=TO;
 RA Hyodo H., Ishikawa Y., Kashiwase K., Ogawa A., Watanabe Y.,
 RA Tsuneyama H., Toyoda C., Uchikawa M., Akaza T., Fujii T.;
 RT "Polymorphisms of RhDya in Japanese."
 RL Submitted (Oct-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB018967; BAA81900.1; -
 DR InterPro: IPR001905; -
 DR InterPro: IPR002229; -
 DR Pfam: PF00909; Ammonium transp; 1.
 DR PRINTS: PR00342; RHESUSRHD.
 DR SEQUENCE 417 AA: 45211 MW: 764B092CBF95ECC4 CRC64;

alignment_scores:

Quality: 2117.00 Length: 417
 Ratio: 5.077 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 99.520

alignment_block:

us-09-600-714-41 x 09UQ22 ..

Align seg 1/1 to: 09UQ22 from: 1 to: 417

1 ATGAGCTCTAAGTACCCGCGGTCTGTCGCGGCTGCCCTGCTGGGC 50
 1 MetSerSerLysTyrProArgSerValAlaArgCysLeuProLeuTrpAl 17
 51 CCTAAGCTGGAAGCAGCTCATCTTCTCTTATTTTTCACCACT 100
 17 aleuThrLeuGlnAlaAlaLeuIleLeuPheTyrPhePheThrHisT 34
 101 ATGACGCTTCTTGAAGATCAAAAGGCGCTGTCGCACTCTATCAAGTT 150
 34 yAspAlaSerLeuGlnAspGlnLysGlyLeuValAlaSerTyrGlnVal 50
 151 GGCCAGATCTGACCGTGAATGCGGCCATTGGCTGGCTTCTCACCTC 200
 51 GlysAlaSerLeuThrValMetAlaAlaIleGlyLeuGlyPheLeuThrSe 67
 201 CAGTTTCCGAGACACAGCTGGAGCAGTGGCTTCAACTCTTCACTGC 250
 67 rSerPheArgArgHisSerTrpSerSerValAlaPheAsnLeuPheMetL 84
 251 TGGGCGTGGTGTGACAGTGGGCAATCCGTCGAGCGGCTTCTGAGCCAG 300
 84 euAlaLeuGlyValGlnTrpAlaIleLeuLeuAspGlyPheLeuSerGln 100
 301 TTCCCTTCTGGGAAGGTGTCATCAACATCTTCACTTTCGCTGGCCAC 350
 101 PheProSerGlyLysValValIleThrLeuPheSerIleArgLeuAlaTh 117
 351 CATGAGTCCTTGTGCGTGCATCTCAGTGCATGCTCTCTGGGGAAG 400
 117 rMetSerAlaLeuSerValLeuIleSerValAspAlaValLeuGlyLysV 134

401 TCAACTTGGCGCAGTTGGTGTGATGGTGTGAGTGACAGCTTTA 450
134 aAsnLeuAlaGlnLeuValValMetValLeuValGluValThralaleu 150
451 GGCACCTGAGGATGTCATCATGATATCTTCAACACAGACTACCAT 500
151 GlysLeuArgMetValIleSerAsnIlePheAsnThrAspTyrHisMe 167
501 GAACATGATGACATTCAGCGTGCAGGCTATTGGCGCTGCTGGG 550
167 CAsMetMetHisIleTyrValPheAlaAlaTyrPheGlyLeuSerValA 184
551 CTTGGTCCCTGCCAAAGCCTTACCCGAGGGAACGAGATGAAGATCAG 600
184 latrPcysLeuProLysProLeuProGluGlyThrGluAspLysAspGln 200
601 ACAGCAACGATACCCAGTTGTGTGTCATGCTGGGCGCCCTTCTGTG 650
201 ThrIatThrIleProSerLeuSerAlaMetLeuGlyAlaLeuPheLeuTr 217
651 GATGTGTGGCGCAGTTTCAACTGCTGCTGCTGAGAAAGTCCATCGAA 700
217 pMePheTrProSerPheAsnSerAlaLeuLeuArgSerProIleGlnA 234
701 GGAAGATGCCGCTGTCACACACTACTATGCTGTAGAGTCAGCGTGGTG 750
234 rGlysAsnAlaMetPheAsnThrTyrTyrAlaValAlaValSerValVal 250
751 ACAGCATCTCAGGGTCATCTTGGCTACCCCGGAGGAAGATCAGCAA 800
251 ThrIatIleSerGlySerSerLeuAlaHisProGlnGlyIleSerly 267
801 GACTATGTGACAGTGGCGGTGGAGGAGGCGTGGCTGGTGGTACCT 850
267 sThrTyrValHisSerAlaValLeuAlaGlyGlyValAlaValGlyThrS 284
851 CGTGTACACTGATCCCTTCCGCTGCTGCTGCCAAGTGTGCTGTGTG 900
284 eRcysHisIleuIleProSerProTrIleuAlaMetValLeuGlyLeuVal 300
901 GCTGGCTGATCTCCGTCGGGGGAGCCAAAGTACCTCCCGGGCTGTGTA 950
301 AlaGlyLeuIleSerValGlyGlyAlaIlySerlyLeuProGlyCysCysAs 317
951 CCGAGTGTGGGAGTCCCGCACAGCTGCATGAGGCTACAACTTCAGCT 1000
317 natrValLeuGlyIleProHisSerIleMetGlyTyrAsnPheSerL 334
1001 TGCTGGGTCTGCTTGAGAGATCATCTACATTGCTGCTGCTGCTGAT 1050
334 euLeuGlyLeuLeuGlyGluIleIleTyrIleValIleLeuValLeuAsp 350
1051 ACCGTGGAGCCGGCAATGCAATGATGGCTTCCAGGCTCCTCAGCAT 1100
351 ThrValGlyAlaGlyAsnGlyMetIleGlyPheGlnValIleuLeuSerI 367
1101 TGGGGAAGCTAGCTGGGCATGATGATGATGATGATGATGATGATGATG 1150
367 eGlyGluLeuSerLeuAlaIleValIleAlaLeuThrSerGlyLeuLeuTr 384
1151 CAGGTTTGTCTCTAAATCTTAAATATGGAAGCACTCATGAGGCTAAA 1200
384 hrgIyLeuLeuLeuAsnLeuLysIleTrpLysAlaProHisGluAlaLys 400
1201 TATTTTGATGACCAAGTTTCTGAGATTTCCTCATTTGCTGTGGATT 1250
401 TyrPheAspArgGlnValAlaPheTrpLysPheTronHisLeuAlaValGlyPh 417
1251 T 1251
417 e 417

seq_name: sp_human:Q9NXY8
seq_documentation block:
ID Q9NXY8 PRELIMINARY; PRT; 417 AA.
AC Q9NXY8;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-MAR-2001 (Tremblrel. 15, Last annotation update)
DE RHESUS BLOOD GROUP, D ANTIGEN.
GN RHD.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Wagner F.F., Flegel W.A.;
RT "Molecular structure of weak D type 18, another RHD allele detected
RT among weak D phenotypes."
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ287289; CAB75731.1; JOINED.
DR EMBL; AJ287291; CAB75731.1; JOINED.
DR EMBL; AJ287292; CAB75731.1; JOINED.
DR EMBL; AJ287293; CAB75731.1; JOINED.
DR EMBL; AJ287294; CAB75731.1; JOINED.
DR EMBL; AJ287295; CAB75731.1; JOINED.
DR EMBL; AJ287296; CAB75731.1; JOINED.
DR EMBL; AJ287297; CAB75731.1; JOINED.
DR EMBL; AJ287298; CAB75731.1; JOINED.
DR InterPro: IPR001905; -
DR InterPro: IPR002229; -
DR Pfam: PF00909; Ammonium transp; 1.
FT PRINTS: PRO0342; RHESUSRHD.
FT VARIANT 1 1 M -> L.
SQ SEQUENCE 417 AA; 45210 MW; 60FB4CF816F163EF CRC64;

alignment_scores:
Quality: 2115.00 Length: 417
Ratio: 5.084 Gaps: 0
Percent Similarity: 99.760 Percent Identity: 99.760

alignment_block:
US-09-600-714-41 x Q9NXY8

Align seg 1/1 to: Q9NXY8 from: 1 to: 417

1 ATGAGCTCTAAGTACCCGCGGTGTCGCGGCTGCTGCGCCCTGTGGC 50
1 MetSerSerLysTyrProTrpSerValAlaArgCysLeuProLeuTrpAl 17
51 CCTAACACTGAGACGAGCTCTCATTTCTCTCTATTTTTCACCACT 100
17 aLeuThrLeuGlnAlaAlaLeuIleLeuLeuPheTyrPhePheThrHisT 34
101 ATGACGCTTCTCTAGAGATCAAAAGGGGCTGCGGCAATCTATCAAGT 150
34 yAspAlaSerLeuGlnAspGlnLysGlyLeuValAlaSerTyrGlnVal 50
151 GGCCAAAGTCTGACCGTGATGGCGCCATTGGCTTGGGCTTCTCACCTC 200
51 GlyGlnAspLeuThrValMetAlaAlaIleGlyLeuGlyPheLeuThrSe 67
201 GAGTTTCCGGAGACACAGCTGAGAGCTGTGGCTTCAACTCTTCATGC 250
67 rSerPheArgArgHisSerTrpSerSerValAlaPheAsnLeuPheMetL 84
251 TGGCGCTGTGTGTCAGTGGGCAATCTGCTGACAGCGCTTCTCAGCCAG 300
84 euAlaLeuGlyValAlaGlnTrpAlaIleLeuLeuAspGlyPheLeuSerGln 100
301 TTCCTTCTGGGAAGCTGTGCATCAGACTTTCAGTATTCGGCTGGCCAC 350

```

|||||
101 pheProserGlyValValIleThrLeuPheSerIleArgLeuAlaTh 117
351 CATGAGTCTTGTGGGTCGATCTCAGTCAGTGGATCTGTGGGAGG 400
117 rMeSerIleuSerValIleuIleSerValAspAlaValLeuGlyLysV 134
401 TCAACTTGGCCAGTTGGTGTGATGTGTGTGTGAGAGTACAGCTTGA 450
134 alaLeuAlaGlnLeuValValMetValLeuValGlnValThrAlaLeu 150
451 GGCAACTGAGGATGTCATCAGTAAATCTTCAACACAGACTACCAT 500
151 GlyAsnLeuArgMetValIleSerAsnIlePheAsnThrAspTyrHisMe 167
501 GAACATGATGACATCTACGTCGCGCCTATTGGCTGCTGTGG 550
167 tAsnMetHisIleTyrValPheAlaAlaTyrPheGlyLeuSerValA 184
551 CCTGGTGCCTCCAAAGCCTTACCCGAGGAGACGAGATTAAGATCAG 600
184 latrPCysLeuProLysProLeuProGlnGlyThrGlnAspLysAspGln 200
601 ACAGCAACGATACCAGTTTGTGTGCATGCTGGGCGCCTTCTGTGTG 650
201 ThrIleThrIleProSerLeuSerAlaMetLeuGlnAlaLeuPheLeuTr 217
651 GATTTTTCGGCAAGTTTCACTCTGCTGCTGTGAGAGTCCATCCGAA 700
217 pMeThrProSerPheAsnSerAlaLeuLeuArgSerProIleGlnA 234
701 GGAAGAAATGCCGTGTCAACACCTACTAGTGTAGCAGTACAGCTGTG 750
234 rGlyAsnAlaValAlaPheSerThrTyrTyrAlaValAlaValSerVal 250
751 ACAGCCATCTCAGGGTCACTCTGGCTACCCCAAGGAGAAATCAGCA 800
251 ThrAlaIleSerGlySerSerLeuAlaHisProGlnGlyLysIleSer 267
801 GACTTATGTGCACAGTGGGCTGTGGCGAGAGGCTGCTGGGTACT 850
267 sThrTyrAlaHisSerAlaValLeuAlaGlyValAlaValAlaGlyHis 284
851 CGTGTACACTGATCCCTTCTCCGTGGCTTGCATGTGCTGGTCTGTG 900
284 erCysHisLeuIleProSerProTrpLeuAlaMetValLeuGlyLeuVal 300
901 GCTGGGCTGATCTCCGTGGGGAGCCAGTACTGCCGGGCTGTGTAA 950
301 AlaGlyLeuIleSerValGlyGlyAlaLysTyrLeuProGlyCysCysAs 317
951 CCGAGTGTGGGAGATCCCCACAGCTCATGGGCTACAACTTCACT 1000
317 nArgValLeuGlyIleProHisSerIleMetClyTyrTrpAsnPheSerL 334
1001 TGTGGGTGTGCTTGGAGAGATCATACATGTGTGCTGTGCTGTGAT 1050
334 euLeuGlyLeuLeuGlyGlnIleIleTyrIleValLeuLeuValLeuAsp 350
1051 ACCGTGCGAGCCGCAATGGCATGTATGGCTTCCAGGCTTCCCTCAGAT 1100
351 ThrValGlyAlaGlnLysAsnGlyMetIleGlyPheGlnValLeuLeuSerI 367
1101 TGGGGAATCAGCTTGGCCATGTCGATAGCTTCCAGCTGTGCTCCGGA 1150
367 eclyGlnLeuSerLeuAlaIleValIleAlaLeuThrSerGlyLeuLeuTr 384
1151 CAGTTTGTCTCTAAATCTTAAATATGGAAGACCTCATGAGGCTAAA 1200
384 hrclyLeuLeuLeuAsnLeuLysIleTrpLysAlaProHisGlyAlaLys 400
1201 TATTTTGTATGACCAAGTTTCTGGAAGTTTCTCATTTGGCTGTGGATT 1250
|||||

```

```

401 TyrPheAspAspGlnValPheTrpLysPheProHisLeuAlaValGlyPh 417
1251 T 1251
417 e 417
seq_name: sp_human:09023

```

```

seq_documentation_block:
ID 09023 PRELIMINARY; PRT; 417 AA.
AC 09023;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, last sequence update)
DT 01-JUN-2000 (TREMBLrel. 14, last annotation update)
DE RH BLOOD GROUP D ANTIGEN (RHD).
GN RHDVA.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCB1_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FK;
RA Hyodo H., Ishikawa Y., Kashiwase K., Ogawa A., Watanabe Y.,
RA Tsuneyama H., Toyoda C., Uchikawa M., Akaza T., Fujii T.;
RT "Polymorphisms of RhDva in Japanese."
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
DR EMBU: AB018966; BAA81899.1; -.
DR InterPro: IPR001905; -.
DR InterPro: IPR002229; -.
DR Pfam: PF00909; Ammonium_transp; 1.
DR PRINTS: PR00342; RHESUSRHD.
SO SEQUENCE 417 AA; 45131 MW; 6D5F4D6CA0DAA884 CRC64;

```

```

alignment_scores:
Quality: 2113.00 Length: 417
Ratio: 5.079 Gaps: 0
Percent Similarity: 99.760 Percent Identity: 99.520

```

```

alignment_block:
US-09-600-714-41 x 09023 ..

```

```

Align seg 1/1 to: 09023 from: 1 to: 417

```

```

1 ATGAGCTCTAAGTACCGCGGTGTGTCGGGCGTCCGCGCTTGGGC 50
|||||
1 MetSerSerLysTyrProArgSerValAlaArgCysLeuProLeuTrpAl 17
51 CCTAACACTGGAGACAGCTCATTCCTCTCTATTTTACCACT 100
|||||
17 aleuThrLeuGlnAlaAlaLeuIleLeuLeuPheTyrPhePheThrIst 34
101 ATGACGCTTCTTAGAGGATCAAAAGGCGCTCGGCATCTATCAGATT 150
|||||
34 yrasPalasSerLeuGlnAspGlnLysGlyLeuValAlaSerTyrGlnVal 50
151 GGCCAAGATCTGACCGTATGAGCGGCGCATGGCTTGGGCTTCTCACCTC 200
|||||
51 GlnGlnAspLeuThrValMetAlaAlaIleGlyLeuGlyPheLeuThrse 67
201 GAGTTTCCGAGACACAGCTGGAGACAGTGGCTTCAACCTTTCATGC 250
|||||
67 rSerPheArgTrpHisSerTrpSerSerValAlaPheAsnLeuPheMetL 84
251 TGGCGCTTGTGTGTCATGAGGCAATCTGCTGAGCGGCTTCTTACAGCAG 300
|||||
84 euAlaLeuGlyValGlnTrpAlaLeuLeuAspGlyPheLeuSerGln 100
301 TTCCCTTCTGGAGAGTGTGCATCACACTGTTCAGTATTCGGCTGGCCAC 350
|||||
101 PheProserGlyLysValValIleThrLeuPheSerIleArgLeuAlaTh 117
|||||

```

```

351 CATGAGTCCTTTGTCGGTCGATCATGAGTGCCTCTTGGGAGG 400
|||||
117 rmetserAlaLeuSerValLeuIIeSerValAspAlaValLeuGlyLysV 134
401 TCACCTTGGGCACTTGGTGTGATGCTGCTGGTGGAGGTACAGCTTGA 450
|||||
134 alaLeuAlaGlnLeuValValMetValLeuValGlnValTThAlaLeu 150
451 GGCACACCTGGAGGATGTCATCATGTAATATCTTCAACACAGACTCCACAT 500
|||||
151 GlnsLeuArgMetValIIeSerAsnIIePheAsnIIePheAspTyrHisWe 167
501 GAACATGATGCACATCTACGTGTTCGACGCTATTTGGGCTGTCTGTGG 550
|||||
167 tAsnMetMetHisIIeTyrValPheAlaAlaTyrPheGlyLeuSerValA 184
551 CCTGTGCTCTCCAAAGCTCTTACCCGAGGAGGAGGAGGATTAAGATCAG 600
|||||
184 latrPcysLeuProLysProLeuProGlnGlyThrGlnAspLysAspGln 200
601 ACAGACATGATACCAAGTTGCTGTCATGCTGGGCGCCCTCTCTGTGTG 650
|||||
201 ThrAlaThrIIeProSerLeuSerAlaMetLeuGlyAlaLeuPheLeuTr 217
651 GATGTTCTGGCCAAAGTTTCAACTGCTGTCTGTCTGAGAAATCCAAATCGAA 700
|||||
217 pMetPheTrpProSerValAsnSerAlaLeuLeuArgSerProIIeGlnA 234
701 GGAAGATGCGCGTGTTCACACACTACTATGCTGTAGCAGTACGCTGTGTG 750
|||||
234 rGlyAsnAlaValAlaPheAsnThrTyrTyrAlaValAlaValSerValA 250
751 ACAGCATCTCAGGGGTATCCCTTGGCTCACCCCAAGGAGGATAGCA 800
|||||
251 ThrAlaIIeSerGlySerSerLeuAlaHisProGlnGlyLysIIeSerLys 267
801 GACTTATGTGCACAGTGCAGTGTGGCAGAGGCGCTGGCTGTGGTACT 850
|||||
267 sThrTyrAlaHisSerAlaValLeuAlaGlyGlyAlaValAlaGlyLys 284
851 CGTGTCACCTGATCCCTTCTCGTGGCTTGCATGCTGTGGCTGTGTG 900
|||||
284 erCysHisLeuIIeProSerProTrieuAlaMetValLeuGlyLeuVal 300
901 GCTGGGCGATCTCCGTCGGGGAGGACCAAGTACCTGCCGGGCTGTGAA 950
|||||
301 AlaGlyLeuIIeSerValGlyGlyAlaLysTyrLeuProGlyCysCysAs 317
951 CCGAGTGTGGGATTTCCCCACAGCTCCATCATGGGCTACAACTTCACCT 1000
|||||
317 nArgValLeuGlyIIeProHisSerSerIIeMetGlyTyrAsnPheserL 334
1001 TGCCTGGTCTGCTTGGAGATCATCTACATTTGCTGCTGCTGCTTGA 1050
|||||
334 euLeuGlyLeuLeuGlyGlnIIeTyrIIeTyrIIeValLeuLeuValLeuAsp 350
1051 ACCGTGGAGCGCGCAATGCGATGATGGCTTCCAGGTCCTCCCTCAGAT 1100
|||||
351 ThrValGlyAlaGlyAsnGlyMetIIeGlyPheGlnValLeuLeuSerII 367
1101 TGGGAACTCAGCTTGGCCATCGTATAGCTCTCAGCTGTGGTCTCCGTA 1150
|||||
367 eGlyIuLeuSerLeuAlaIIeValIIeAlaLeuThrSerGlyLeuLeuT 384
1151 CAGGTTCTCTCTAAATCTTAAATATGGAACACCTCATGAGGCTTAAA 1200
|||||
384 nGlyIuLeuLeuAsnLeuLysIIeTrrPlyAlaProHisGlnAlaLys 400
1201 TATTTGATGACCAAGTTTCTGGAAGTTTCTCATTTGGCTGTGGTGA 1250
|||||
401 TyrPheAspAspGlnValPheTrpLysPheProHisLeuAlaValGlyPh 417
1251 T 1251

```

```

seq_name: sp_human:Q9UQ21
417 e 417
seq_documentation_block:
ID Q9UQ21 PRELIMINARY; PRT; 417 AA.
AC Q9UQ21;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-JUN-2000 (Tremblrel. 14, Last annotation update)
DE RH BLOOD GROUP D ANTIGEN (RHD).
GN RHDVA.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBL_TaxID=9606;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=TT;
RA Hyodo H., Ishikawa Y., Kashiwase K., Ogawa A., Watanabe Y.,
RA Tsuneyama H., Toyoda C., Uchikawa M., Akaza T., Fujii T.;
RT "Polymorphisms of Rhdva in Japanese.";
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB018968; BAA81901.1; -
DR InterPro; IPR001905; -
DR InterPro; IPR002229; -
DR Pfam; PF00909; Ammonium transp; 1.
DR PRINTS; PR00342; RHESUSRD.
SQ SEQUENCE 417 AA; 45177 MW; B7327B16747E61D CRC64;

alignment_scores:
Quality: 2107.00 Length: 417
Ratio: 5.065 Gaps: 0
Percent Similarity: 99.760 Percent Identity: 99.041

alignment_block:
US-09-600-714-41 x Q9UQ21 ...

Align seg 1/1 . to: Q9UQ21 from: 1 to: 417

1 ATGAGCTCTAAGTACCCGCGGTCTGTCGGCGCTGCGCCCTCTGGGC 50
|||||
1 MetSerSerLysTyrProArgSerValArgArgCysLeuProLeuTrrpAl 17
51 CCTAACACTGGAGACGCTTCATTTCTCTCTCTATTTTTCACCACT 100
|||||
17 alaThrLeuGlnAlaAlaAlaLeuIIeLeuLeuPheTyrPhePheThrHisT 34
101 ATGAGCTTCTTAGAGGATCAAAAGGGCTGCGTGGCATCCATCAAGTT 150
|||||
34 yrrAspAlaSerLeuGlnAspGlnLysGlyLeuValAlaSerTyrGlnVal 50
151 GGCCAAAGATCTGACCGTATGGCGGCGCATTTGGCTTGGCTTCTCACCTC 200
|||||
51 GlynAsnAspLeuThrValMetAlaAlaIIeGlyLeuGlyPheLeuThrSe 67
201 GAGTTTCCGAGACACACAGCTGGAGCAGTGTGGCTTCAACCTTTCAATC 250
|||||
67 rSerPheArgArgHisSerTrrPserSerValAlaPheAsnLeuPheMetL 84
251 TGGCGCTGTGTGAGGAGGCAATCCGTCGGAGCGCTTCTCGAGCGAG 300
|||||
84 euAlaLeuGlyValGlnTrrpAlaIIeLeuLeuAspGlyPheLeuSerGln 100
301 TTCCCTTGGGGAAGGTGTCATCACACTGTTCATATTTGGCTGGCCAC 350
|||||
101 PheProSerGlyLysValValIIeThrLeuPheSerIIeArgLeuAlaTh 117
351 CATGAGTGTCTTGTGGTGTGATCTCAGTGGATGCTGTCTTGGGGAAG 400
|||||
117 rMetSerAlaLeuSerValLeuIIeSerValAspAlaValLeuGlyLysV 134

```

```

401 TCAACTGGCCAGTGTGGTGGATGTCGTGGAGGTGACAGCTTTA 450
134 aLaSnLeuAlaGlnLeuValValMetValLeuValGlnValThraLeu 150
451 GGCAACCTGAGAGATGTCATGATATATCTTCAACACAGCATCCACAT 500
151 GlySnLeuArGmetValIleSerAsnIlePheAsnThrAspTyrHisLe 167
501 GAACATGATGACATCTAGCTGTCCGAGCCTATTGGCTGCTGTGG 550
167 CasMetHeChisIleTyrValPheAlaIleTyrPheGlyLeuSerValA 184
551 CCTGGTCCCTCCAAAGCCTCTACCCGAGGACGAGAGATAAAGATCAG 600
184 IaTrpCysLeuProIlySerProIleuProGlnGlyThrGlnAspLys 200
601 ACAGCAGATACCCAGTTGTCTGCGCATGCTGGCGCCCTCTTCTGTG 650
201 ThraIaThrIleProSerLeuSerAlaMetLeuGlnAlaLeuPheLeu 217
651 GATGTTGGCCAGTTTCAACTGCTGCTGCTGAGAGTCCATCCGAA 700
217 PhePheTrpProSerValAsnSerAlaLeuLeuArgSerProIleGlnA 234
701 GGAAGAAATGCCGTGTTCACACCTACTAGCTGTAGCAGTCAGCGTGTG 750
224 rGlySAsnAlaMetPheAsnThrTyrTyrAlaLeuAlaValSerValVal 250
751 ACAGCAGATCTAGGGTCTCTGGCTGACCCCAAGGAAAGATCAGCA 800
251 ThrAlaIleSerGlySerSerLeuAlaHisProGlnGlyLysIleSer 267
801 GACTATGTGCACAGTGGGTGTGGCAGAGGCGTGGCTGGTACT 850
267 sThrTyrValHisSerAlaValLeuAlaGlyGlyValAlaValGlyThr 284
851 CGTGCACCTGATCCCTCTCCGCTGCTGCGCATGCTGCTGGTCTGTG 900
284 eRcYsHisLeuIleProSerProTrpLeuAlaMetValLeuGlyLeuVal 300
901 GCTGGCGTGCATCCGTCGGGGGAGCCAGTACTCCCGGGGTGTGTA 950
301 AlaGlyLeuIleSerValGlyGlyAlaLysTyrLeuProGlyCysCysAs 317
951 CCGAGTGTGGGATCCCCACAGTCCATGAGGCTACAACTTCACT 1000
317 nArgValLeuGlyIleProHisSerSerIleMetGlyTyrAsnPheSer 334
1001 TGCTGGTCTGCTGGAGAGATCATCAATGTGCTGCTGGTCTTGAT 1050
334 euLeuGlyLeuLeuGlyGlnIleIleTyrIleValLeuLeuValLeuAsp 350
1051 ACCGTGCGAGCCGCAATGGCATGATTTGGCTCCAGTCCCTCCAGAT 1100
351 ThrValGlyAlaGlyAsnGlyMetIleGlyPheGlnValLeuLeuSer 367
1101 TGGGAACTAGCTTGGCCATGATAGTCTGACAGTCTGCTGCTCGA 1150
367 eGlyGlnLeuSerLeuAlaIleValIleAlaLeuThrSerGlyLeuLeu 384
1151 CAGGTTTGCTCTAAATCTTAAATATGGAAGAAGCCTCATGAGGCTAA 1200
384 nrgGlyLeuLeuLeuAsnLeuLysIleTrpLysAlaProHisGlnAlaLys 400
1201 TATTTGTATGACCAAGTTTCTGGAAGTTTCTCATTTGGCTGTGGATT 1250
401 TyrPheAspAspGlnValPheTrpLysPheProHisLeuAlaValGlyPh 417
1251 T 1251
417 e 417

```

```

seq_name: sp_human:099906
seq_documentation_block:
ID 099906 PRELIMINARY; PRT; 417 AA.
AC 099906;
DT 01-MAY-1997 (TREMBLrel. 03, Created)
DT 01-MAY-1997 (TREMBLrel. 03, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE D CATEGORY IIIC ANTIGEN.
GN RHD.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN (1)
RP SEQUENCE FROM N.A.
RA MEDLINE=96269518; PubMed=8669091;
RA Beckers E.A., Faas B.H., Lijthart P., Simsek S., Overbeek M.A.,
RA dem Borne A.E., van Rhemen D.J., der Schoot C.E.;
RT "Characterization of the hybrid RHD gene leading to the partial D
RT category IIIC phenotype.";
RL Transfusion 36:567-574(1996).
DR EMBL: S82449; AAB37636.2; -.
DR InterPro: IPR001905; -.
DR InterPro: IPR002229; -.
DR Pfam: PF00909; Ammonium_transp; 1.
DR PRINTS: PR00342; RHESUSRHD.
SQ SEQUENCE 417 AA; 45099 MW; F2A1ECCB5B921260 CRC64;

```

```

alignment_scores:
Quality: 2104.00 Length: 417
Ratio: 5.058 Gaps: 0
Percent Similarity: 99.760 Percent Identity: 99.041

```

```

alignment_block:
US-09-600-714-41 x 099906 ..

```

```

Align seg 1/1 to: 099906 from: 1 to: 417

```

```

1 ATGAGCTAGTACCGCGGCTGTCCGGCGCTGCCCTCTGAGGC 50
|||||
1 MetSerSerLysTyrProArgSerValArgArgCysLeuProLeuTrpAl 17
51 CCTAACACTGGAACACACTCTCATCTCTCTATTTTTCACCCACT 100
|||||
17 AleuThrLeuGlnAlaAlaLeuIleLeuLeuPheTyrPhePheThrHis 34
101 ATGACGCTTCCTTAGAGATCAAAAGGCGCTGTGSCATCCTATCAAGTT 150
|||||
34 YrAspAlaSerLeuGlnAspGlnLysGlyLeuValAlaIleSerTyrGlnVal 50
151 GGCCAAATGTGACCGTGAATGGCGGCATGGCTGGGCTTCCTCACCTC 200
|||||
51 GlyGlnAspLeuThrValMetAlaAlaIleGlyLeuGlyPheLeuThrSe 67
67 rSerPheArgArgHisSerTrpSerSerValAlaPheAsnLeuPheMet 84
251 TGGCGCTTGGTGTGCAGTGGGCAATCTGCTGACAGGCTTCCTGAGCCAG 300
|||||
84 euAlaLeuGlyValGlnTrpAlaIleLeuLeuAspGlyPheLeuSerGln 100
301 TTCCCTCTGGGAGATGGTGCATACACATGTCATTTGGCGTGGCCAC 350
|||||
101 PheProSerGlyLysValValIleThrLeuPheSerIleArgLeuAlaTh 117
351 CATGAGTGTCTTGCCTGATCTCAGTGCATGATGCTGTCTTGGGGAAG 400
|||||
117 rMetSerAlaMetSerValLeuIleSerAlaGlyAlaValLeuGlyLysV 134
401 TCAACTTGGCGCAGTTGTGGTGGATGCTGCTGCTGAGGTGACAGCTTTA 450

```

[illegible]

```

seq_documentation_block:
ID 09UK19 PRELIMINARY: PRT: 417 AA.
AC 09UK19;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)
DE RHD TYPE IIIA PROTEIN.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
OX NCBI_TaxID=9606;
[1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97398395; PubMed=9256293;
RA Huang C.H., Chen Y., Reid M.;
RT "Human DcIta1 erythrocytes: Rhd protein is associated with multiple
RL Am. J. Hematol. 55:139-145(1997).
[2]
RP SEQUENCE FROM N.A.
RA Huang C.H.;
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF187846; AA04565.1; -.
DR InterPro; IPR001905; -.
DR InterPro; IPR002229; -.
DR Pfam; PF00909; Ammonium transp. 1.
DR PRINTS; PR00342; RHESUSRHD.
SQ SEQUENCE 417 AA: 4523 MW: 491FB1FE1BA37957 CRC64;

```

alignment_scores:	Quality: 2094.00	Length: 417
Ratio: 5.058	Gaps: 0	
Percent Similarity: 99.281	Percent Identity: 98.801	
alignment_block:		
US-09-600-714-41 x Q9UK19 ..		
Align seg 1/1	to: Q9UK19	from: 1 to: 417
1	ATGAGCTCTAAGTACCCGGCGTCTGTCCGGGCGTCCGCCCCCTGTGGC	50
1	MetSerSerIysTrpProArgSerValArgArgCysLeuProLeuTrpAl	17
51	CCTAACACTGGAACAGCTCTCATTTCTCTCTTCAATTTTACCACCT	100
17	aleuThrLeuGlnAlaIleuLeuIleLeuLeuPheTrpPhePheTrhHis	34
101	ATGACGGTCTCCTTGAGAGATCAAAAGGGGCGTCGTGGACCTATCAAGTT	150
34	yrAspAlaSerLeuGlnAspGlnLysGlyLeuValAlaLaserTrglnVal	50
151	GGCCAGATCTGACCGGTGATGGCGGCATTGGCTTGGGCTTCACGTC	200
51	GlyGlnAspLeuThrValMetAlaIleuGlyLeuGlyPheLeuThrSe	67
201	GAGTTTCCGGAGAACACGCTGGAGAGTGGGCTTCAACCTTCATATG	250
67	rIserPheArgTrhHisSerTrpSerSerValAlaPheAsnLeuPheMetL	84
251	TGGCGCTTGGTGATCGACAGTGGGCAATCTGCTGGAGCGGCTTCCGTGACCG	300
84	eulAlaLeuGlyValGlnTrpAlaIleLeuLeuAspGlyPheLeuSerGln	100
301	TTCCCTTCTGGGAAGGTGGTCATCACACTGTTCACTATTCCGCTGGCCAC	350
101	PheProSerGlyLysValIleAlaIleThrLeuPheSerIleArgLeuAlaTh	117
351	CATGAGAGCTTTGGCGGCTGATCTTCAGTGGATGCTGCTTGGGGAGAG	400
117	rMetSerAlaLeuSerValIleLeuIleSerValAspAlaValIleuGlyLysV	134
401	TCAACTTGGCGCACTTGGTGTGATGTGCTGGTGGAGGTGACAGCTTTA	450

```

|||||
134 aLaSnLeuAlaGlnLeuValValMetValLeuValGluValThrAlaLeu 150
451 GGCAACCTGAGGATGTCATCATATATCTTCAACACAGACTACACAT 500
151 GlnThrLeuArgMetValIleSerAsnIlePheAsnThrAspTyrHisMe 167
501 GAACATGATGACATCTACGTGTTCGACGCTATTTTGGCTGCTGTG 550
167 tAsnMetethIstIleTyrValAlaPheAlaIleTyrPheGlnLeuSerValA 184
551 CCTGTGCTGCTCCAAAGCCTTACCCGAGGACGAGAGATAAAGATCAG 600
184 latrPcysLeuProLysProLeuProGlnGlnTyrGlnAspLysAspGln 200
601 ACAGCAACGATACCAAGTTGTGTCGATGTCGGGCGGCTCTCTGTG 650
201 ArgAlaThrIleProSerLeuSerAlaMetLeuGlnAlaLeuPheLeuT 217
651 GATGTCTTGCCCAAGTTTCAACTGCTGTGCTGAGAAATCAATCGAA 700
217 pMetPheThrProSerValAlaSerAlaLeuLeuArgSerProIleGlnA 234
701 GGAAGATGCCGTGTTCACACACTACTAGTGTGACAGTACGCTGTG 750
234 rGlyAsnAlaValAlaPheAsnThrTyrTyrAlaValAlaValSerValVal 250
751 ACAGGCATCTGAGGTCATCTGTCGTCACGCCCAAGGAGATCAGCA 800
251 ThrAlaIleSerGlnSerSerLeuAlaHisProGlnArgLysIleSerL 267
801 GACTATGTGCACAGTGCCTGTTCGACGAGGCGTGGCTGTGGATACCT 850
267 sThrTyrValHisSerAlaValLeuAlaGlnGlyValAlaValGlnTyr 284
851 CGTGCACCTGATCCCTTCCGTGGCTGTCGCAATGGCTGCTGTGTG 900
284 eRcysHisLeuIleProSerProTrpLeuAlaMetValLeuGlnLeuVal 300
901 GCTGGGCTGATCTCGTCGGGGGAGCCAAATGACCTGGCGGCTGTGAA 950
301 AlaGlnLeuIleSerValGlnGlyAlaLysTyrLeuProGlnCysCysAs 317
951 CCGAGTGTGGGATTCGCCACAGCTCCATCATGGGCTTACACTTCAGCT 1000
317 nArgValLeuGlnTyrIleProHisSerSerIleMetGlnTyrAsnPheSerL 334
1001 TGTGGGCTCTCTGGAGAGATCATATGATTTGCTGCTGCTGTGAT 1050
334 euleuGlnLeuLeuGlnGlnIleIleTyrIleValLeuLeuValLeuAsp 350
1051 ACCGTCGAGCGCGCAATGGCATGTTGGCTTCGAGGCTCCCTCAGAT 1100
351 ThrValGlnAlaGlnLysnGlnMetGlnGlnPheGlnValLeuLeuSerL 367
1101 TGGGAACTCAGCTTGGCCATCGATAGCTGTCTCAGCTGTGCTCTGA 1150
367 eGlnGlnLeuSerLeuAlaIleValIleAlaLeuThrSerGlnLeuLeuT 384
1151 CAGGTTGCTCTTAATCTTAATATGGAAGAAGCACTCATGAGGCTTAA 1200
384 hrGlnLeuLeuLeuAsnLeuLysIleTyrLysAlaProHisGlnAlaLys 400
1201 TATTTGATGACCAAGTTTCTGGAAGTTTCTCATTTGGCTGTGGAT 1250
401 TyrPheAspAspGlnValAlaPheThrLysPheProHisLeuAlaValGln 417
1251 T 1251
417 e 417
seq_name: sp_human:O9HCC3

```

```

seq_documentation_block:
ID O9HCC3 PRELIMINARY; PRT; 417 AA.
AC O9HCC3:
DT 01-MAR-2001 (TReMBLrel. 16, Created)
DT 01-MAR-2001 (TReMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TReMBLrel. 16, Last annotation update)
DE RH BLOOD D GROUP ANTIGEN POLYPEPTIDE.
GN RHD.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Kashiwase K., Uchikawa M., Ishikawa Y.;
RT "Rh blood D group antigen polypeptide, RHDK.";
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB049754; BAB16598.1; -;
SQ SEQUENCE 417 AA; 45136 MW; 2CA89E3E716087C8 CRC64;

alignment_scores:
Quality: 2094.00 Length: 417
Ratio: 5.046 Gaps: 0
Percent Similarity: 99.520 Percent Identity: 98.321

alignment_block:
US-09-600-714-41 x O9HCC3
Align seg 1/1 to: O9HCC3 from: 1 to: 417

1 ATGAGCTCTAAGTACCGCGTCTGTCCGGCGCTCCGCGCTCTGTGGC 50
1 MetSerSerLysTyrProArgSerValArgArgCysLeuProLeuThrAl 17
51 CCTTACACTGGAGCAGCTCTCATTTCTCTTATTTTAACTTACCAC 100
17 aleuThrLeuGlnAlaAlaLeuIleLeuLeuPheTyrPhePheThrIst 34
101 ATGAGGCTTCCTTGTAGAGATCAAAAGGGGCTCGTGGCATCTTACAG 150
34 yrasPheAsnLeuLeuGlnAspGlnLysGlnLeuValAlaSerTyrGlnVal 50
151 GGCCAAATCTGACCTGATGGCGGCAATGGCTTGGGCTTCTCCACTC 200
51 GlnGlnAspLeuThrValMetAlaAlaLeuGlnGlnLeuThrLeuThr 67
201 GAGTTTCCGGAGACACAGCTGGAGCAGTGTGGCTTCAACTCTTATGC 250
67 rAsnPheArgThrHisSerTrpSerSerValAlaPheAsnLeuPheMetL 84
251 TGGGCGCTTGGTGTGACATGGGCAATCTGCTGGAGGCTTCCTGAGCAG 300
84 euAlaLeuGlnValGlnThrPheAlaLeuLeuAspGlnPheLeuSerGln 300
301 TTCCCTTCTGGGAAGTGTGCATCACACTGTTCAGTATTCGGCTGGCC 350
101 PheThrProGlnGlnValValIleThrLeuPheSerIleThrGlnLeuAla 117
351 CATGAGTCTTTTGTGCTGTGATCTCAGTGGATGCTGTCTTGGGGAAG 400
117 rMetSerAlaMetSerValLeuIleSerAlaGlnAlaValLeuGlnLys 134
401 TCACTTGGCGGCACTTGGTGTGATGTGCTGTGGAGGTCGACAGCTT 450
134 aLaSnLeuAlaGlnLeuValValMetValLeuValGlnValThrAlaLeu 150
451 GGCAACCTGAGGATGTCATCATATATCTTCAACACAGACTACACAT 500
151 GlnThrLeuArgMetValIleSerAsnIlePheAsnThrAspTyrHisMe 167
501 GAACATGATGACATCTACGTGTTCGACGCTATTTTGGCTGCTGTG 550

```

```

167  tAsnMetHethisIleTyValPheAlaAlaTyPheGlyLeuSerValA 184
551  CCGTGCTGCTGCAAAAGCCTACCCGAGGAGGAGGATAAAGATCAG 600
184  lAtPrCySLeuProLySProLeuProGlnGlyThrGlnAspLySAspGln 200
601  ACAGCAACGATACCCAGTTTGTCTGCGCATGCTGGGCGCCCTCTCTGTG 650
201  ThrAlaThrIleProSerLeuSerIleMetLeuGlyAlaLeuPheLeuTr 217
651  GATGTTCTGGCCCAAGTTTCAACTGCTGCTGCTGCTGAGAGTCCAAATGAAA 700
217  pMetHethTrProSerPheAsnSerAlaLeuLeuArgSerProIleGlnA 234
701  GGAACAATGCGCGTGTCAACACCTACTGCTGCTGCTGCTGCTGCTGCTG 750
234  tGlySAsnAlaValPheAsnThrTyTyTyAlaValAlaValSerValAl 250
751  ACAGCAATCTAGAGGTGATCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 800
251  ThrAlaIleSerGlySerSerLeuAlaHisProGlnGlyLySleSerLy 267
801  GACTTATGTGCACAGTGGCGGTGGTGGCAGAGGCGTGGCTGGTGGTGGT 850
267  sThrTyTyAlaHisSerAlaValAlaLeuAlaGlyAlaValAlaGlyThrS 284
851  CGGTGCACCTGATCCCTCTGCGGTGGTGGCAGTGGTGGTGGTGGTGGT 900
284  eCySHisLeuIleProSerProIlePheAlaMetValLeuGlyLeuVal 300
901  GCTGGGCTGATCTCCGTCGGGGGAGCCAAATGACTGCGGGGCTGTTGTA 950
301  AlagIleuIleSerValGlyAlaTyTyTyLeuProGlyCySAsp 317
951  CCGAGTGTGGGGGATCCCGACAGCTCCATCATGGGCTCAACAATTCAGCT 1000
317  nArgValLeuGlyIleProHisSerSerIleMetGlyTyTyAsnPheSerL 334
1001  TGCTGGGTGCTGTGGAGAGATCATCTACATGTGCTGCTGCTGCTGCTGAT 1050
334  euLeuGlyLeuLeuGlyGlnIleIleTyTyIleValLeuLeuValLeuAsp 350
1051  ACCGTGGAGCCGCAATGGCATGATGGCTTCCAGTCCCTCCATCAGCAT 1100
351  ThrValGlyAlaGlyAsnGlyMetIleGlyPheGlnValLeuLeuSerL 367
1101  TGGGGAATCAGCTTGGCCATGCTGATAGCTTCAAGTGTGCTGCTGCTGA 1150
367  eGlyLeuLeuSerLeuAlaIleValIleAlaLeuThrSerGlyLeuLeuT 384
1151  CAGGTTTGCTCTTAATCTTAAATATGGAAGAAGCACTCATGAGGCTAAA 1200
384  hArgIleuLeuLeuAsnLeuLySleTrPlyAlaProHisGlnAlaLyS 400
1201  TATTTTGATGACCAAGTTTCTGGAAGTTTCCATTTGGCTGTTGGATT 1250
401  TyTyPheAspAspGlnValPheTrPlySPhProHisLeuAlaValGlyPh 417
1251  T 1251
417  e 417

```

seq_name: sp_human:Q9H248

seq_documentation_block:

ID_Q9H248 PRELIMINARY: PRT: 417 AA.

AC_Q9H248:

DT_01-MAR-2001 (TREMBLERL.16, Created)

DT_01-MAR-2001 (TREMBLERL.16, Last sequence update)

DT_01-MAR-2001 (TREMBLERL.16, Last annotation update)

DE_RHESUS D CATEGORY VI TYPE IV PROTEIN.

GN_RHD.

OS_Homo sapiens (Human).

```

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Esteban R., Nogués N., Montero R., Hurtado M., Maroto S., Ribera A.;
RT "Characterization of a new RHD category VI (type IV) allele,
RT predominant in the Spanish population.";
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF12679; AAC40872.1; -; 03FDD058960E41D89 CRC64;
SQ SEQUENCE 417 AA; 45293 MW; 03FDD058960E41D89 CRC64;

```

```

alignment_scores:
    Quality: 2043.00      Length: 417
    Ratio: 4.971         Gaps: 0
    Percent Similarity: 98.561      Percent Identity: 95.923

```

alignment_block:

Align seg 1/1 to: Q9H248 from: 1 to: 417

```

1  ATGAGCTTCTAGTACCGCGGCTGTGTCGCGCGTGCCTGCCCTCTGGGC 50
   |||||||
1  MetSerSerLySProArgSerValArgGlyLeuProLeuTrpAl 17
51  CCTACACTGAGACAGCTTCATCTCCCTCTCATTTTATTTTACCACCT 100
   |||||||
17  AleuThrLeuGlnAlaAlaLeuIleLeuLeuPheTyTyPhePheThrHisT 34
101  ATGAGCTTCTCTTAGAGATCAAAAGGGGCTGTCGATCCATCAAGTT 150
   |||||||
34  YrAspAlaSerLeuGlnAspGlnLySgLyLeuValAlaSerTyTyGlnAl 50
151  GGCCAGATCTGACCGGTGATGGCGGCATTGGCTTGGCTTCCATCAGCTC 200
   |||||||
51  GlyGlnAspLeuThrValMetAlaAlaIleGlyLeuGlyPheLeuThrSe 67
201  GAGTTTCCGGAGACACAGCTGAGAGAGTGGCTTCAACCTTCATATGC 250
   |||||||
67  rSerPheArgArgHisSerTrpSerSerValAlaPheAsnLeuPheMetL 84
251  TGGCGCTTGTGTGTCAGTGGGCAATCTGCTGAGCGCTTCCGAGCCAG 300
   |||||||
84  euAlaLeuGlyValGlnTrpAlaIleLeuLeuAspGlyPheLeuSerGln 100
301  TTCCCTTGTGGAGGTGTCATCACACTGTCAGTATTGGCTGGCCAC 350
   |||||||
101  PheProSerGlyLySValValIleThrLeuPheSerIleArgLeuAlaTh 117
351  CATGAGTCTGTTGTGGGTGCTGATCTCAAGTGGATGCTGCTGGGGAAG 400
   |||||||
117  rMetSerAlaMetSerValLeuIleSerAlaGlyAlaValLeuGlyLySV 134
401  TCACCTTGGCGCAGTTGTGTGATGTGCTGCTGGTGGAGTGCACACTTA 450
   |||||||
134  AlAsnLeuAlaGlnLeuValValMetValLeuValGlnValThrAlaLeu 150
451  GGCAACCTGAGAGATGTCATCAGTAAATATCTTCAACACAGACTACCACAT 500
   |||||||
151  GlyThrLeuArgMetValIleSerAsnIlePheAsnThrAspTyTyHisMe 167
501  GAACATGATGACATCTAGCTGTTGGAGCCCAATTTTGGGCTGCTGGG 550
   |||||||
167  tAsnLeuArgHisPheTyTyValPheAlaAlaTyTyPheGlyLeuThrValA 184
551  CCGTGCTGCTGCAAAAGCCTACCCGAGGAGGAGGAGATAAAGATCAG 600
   |||||||
184  lAtPrCySLeuProLySProLeuProLySgLyThrGlnAspAsnAspGln 200
601  ACAGCAACGATACCCAGTTTGTCTGCCATGCTGGGCGCCCTTCTTGTG 650
   |||||||

```

```

201 ArgAlaThrIleProSerLeuSerAlaMetLeuGlyAlaLeuPheLeuTr 217
651 GATGTTTCGGCCAGTTTCACCTGCTGCTGCTGAGAACTGCATTCGAAA 700
|||||
217 PhePheTrpProSerValAsnSerAlaLeuLeuArgSerProIleGln 234
|||||
701 GGAAGAAATGCGGTGTTCACACCTACTAGCTGTAGCAGTACGCGTGTG 750
|||||
234 rGlyAsnAlaMetPheAsnThrTyTrAlaLeuAlaValSerValVal 250
|||||
751 ACACCCATCTCAGGGTCACTCCTGGCTCACCCCAAGGAGATCAGCAA 800
|||||
251 ThrAlaIleSerGlySerSerLeuAlaHisProGlnArgLysIleSerMe 267
|||||
801 GACTTATGTGCACAGTCCGCTGTTGGCAGAGCGTGGCTGGTACT 850
|||||
267 tThrTyTrAlaHisSerAlaValLeuAlaGlyAlaValAlaGlyThrS 284
|||||
851 CGTGTACCTGATCCCTTCTCCGTGGCTTGGCATGGTGGCTGTGTG 900
|||||
284 eRcYsHisLeuIleProSerProTrpLeuAlaMetValLeuGlyLeuVal 300
|||||
901 GCTGGCGCTGATCCCTGCGGGGAGCCAACTACCTGCCGGGGTGTGTA 950
|||||
301 AlaGlyLeuIleSerValGlyGlyAlaLysTyLeuProGlyCysCysAs 317
|||||
951 CCGAGTGTGGGGATCCCGACAGCTCATGAGGCTACGAGCTACACTCAGCT 1000
|||||
317 nArgValLeuGlyIleProHisSerSerIleMetGlyTyTrAsnPheSerL 334
|||||
1001 TGCCTGGCTCTCTGGAGAGATCATCTACATTTGCTGCTGCTGTGAT 1050
|||||
334 euLeuGlyLeuLeuGlyIleIleIleTyTrIleValLeuLeuValLeuAsp 350
|||||
1051 ACCGTGCGAGCCGCAATGCGCATGTTGGCTTCCAGGCTCTCCACAGAT 1100
|||||
351 ThrValGlyAlaGlyAsnGlyMetIleGlyPheGlnValLeuLeuSerL 367
|||||
1101 TGGGGAATCAGCTTGGCCATGATAGTACGTCTCAGCTGCTGCTCCCTGA 1150
|||||
367 eGlyGlnLeuSerLeuAlaIleValIleAlaLeuThrSerGlyLeuLeuTr 384
|||||
1151 CAGGTTTGCCTCTAAATCTTAAATATGAAAGACCTCATGAGGCTTAA 1200
|||||
384 hRgLyLeuLeuLeuAsnLeuLysIleTrpLysAlaProHisGlyAlaLys 400
|||||
1201 TATTTTGATGACCAAGTTTCTGGAAGTTTCTCATTTGGCTGTGGAT 1250
|||||
401 TyrPheAspAspGlnValPheTrpLysPheProHisLeuAlaValGlyPh 417
|||||
1251 T 1251
|
417 e 417

```

seq_name: sp_human:Q9NPX0

seq_documentation_block:

ID Q9NPX0 PRELIMINARY: PRT: 417 AA.

AC Q9NPX0:

DT 01-OCT-2000 (Tremblrel. 15, Created)

DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)

DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)

DE RH BLOOD GROUP ANTIGEN RHD.

GN RHIVB(J).

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RA Hyodo H., Ishikawa Y., Tsuneyama H., Kashiwase K., Toyoda C.,

RA Uchiyama M., Akaza T., Fujii T., Kozuma S., Takekani Y., Jui T.,

RT "New Rhivb identified in Japanese."

```

RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB037270; BAA90298.1; -.
DR InterPro: IPR001905; -.
DR InterPro: IPR002229; -.
DR Pfam: PF00909; Ammonium_transp; 1.
DR PRINTS: PR00342; RHESUSRHD.
SQ SEQUENCE 417 AA; 45429 MW; 647D4A4678CB498 CRC64;

```

alignment_scores:

Quality: 2039.00 Length: 417
Ratio: 5.022 Gaps: 0
Percent Similarity: 57.362 Percent Identity: 97.122

alignment_block:

US-09-600-714-41 x Q9NPX0

Align seg 1/1 to: Q9NPX0 from: 1 to: 417

```

1 ATGAGCTCTAAGTACCGCGGCTGTCCGGCGCTGCCCTGGCCCTGGGC 50
|||||
1 MetSerSerLysTyTrProArgSerValArgArgCysLeuProLeuTrpAl 17
|||||
51 CCTAACACTGGAGACAGCTCTCATTTCTCTTCTATTATTTTACCACACT 100
|||||
17 AleuThrLeuGlnAlaAlaLeuIleLeuPheTyTrPhePheThrHisT 34
|||||
101 ATGAGCGCTTCCTTAGAGATCAAAAGGGGCTCGTGGCATCCTATACAGTT 150
|||||
34 YrAspAlaSerLeuGlnAspGlnLysGlyLeuValAlaSerTyGlnVal 50
|||||
151 GGCCAGATCTGACCGTATGAGCGCCATTGGCTTGGGCTTCCACCTC 200
|||||
51 GlyLinspleuThrValMetAlaAlaIleGlyLeuGlyPheLeuThrSe 67
|||||
201 GAGTTTCCGAGACACACAGCTGAGCAGTGGCTTCAACCTTTCATGC 250
|||||
67 rSerPheArgTrpHisSerTrpSerSerValAlaPheAsnLeuPheMetL 84
|||||
251 TGGGCGTTGGTGTGACATGCGGCATCCTGCTGAGCGGCTTCCAGGCAG 300
|||||
84 euAlaLeuGlyValGlnTrpAlaIleLeuLeuAspGlyPheLeuSerGln 100
|||||
301 TTCCCTTCTGGGAGAGTGTGCATCACACTGTTCAATTCGGCTGGCCAC 350
|||||
101 PheProSerGlyLysValValIleThrLeuPheSerIleThrLeuAlaTh 117
|||||
351 CATGAGTCTTGTGCTGCTGATCTCAGTGGATGCTGTCTTGGGGAAAG 400
|||||
117 rMetSerAlaLeuSerValLeuIleSerValAspAlaValLeuGlyLysV 134
|||||
401 TCAACTTGGCGCAGTGTGGTGTATGGTGGCTGGAGGTACAGCTTTA 450
|||||
134 AlAsnLeuAlaGlnLeuValValaMetValaLeuValGlnValThrAlaLeu 150
|||||
451 GGCACTGAGAGATGTCATCATGTAATATCTTCAACACAGACTACCAT 500
|||||
151 GlyAsnLeuArgMetValIleSerAsnIlePheAsnThrAspTyTrHisMe 167
|||||
501 GAACTATGATGACATCTTACGTTGCGCAGCCTATTATTTGGGCTGTG 550
|||||
167 tAsnMetMetHisIleTyTrValPheAlaAlaTyTrPheGlyLeuSerVala 184
|||||
551 CCTGGTGGCTGCCAAGCCTCTACCGGAGGAAAGGAGATTAAGATCAG 600
|||||
184 lArgPcysLeuProLysProLeuProGlnGlyTrpGlnAspLysAspGln 200
|||||
601 ACACCAAGATACCACTTGTCTGCCATGCTGGGCGCCCTTCTTGTG 650
|||||
201 ThrIleThrIleProSerLeuSerAlaMetLeuGlyAlaLeuPheLeuTr 217
|||||
651 GATTTTCTGCGCAAGTTTCAACTGCTGCTGCTGAGAGTCAATTCGAAA 700
|||||

```



```

217 pNcPheTTrProSerPheAsnSerAlaLeuLyrSerProIleGluA 234
701 GGAAGAATGCGCTGTTCACACCTACTATGCTGACAGTCAGCGTGC 750
|||||
234 rGLyAsnAlaValAlpheAsnThrTYrTAlaValAlaValSerValAl 250
751 ACAGGCATCTCAGGGTCATCTCTGGCTCACCCCAAGGAGAGATGACA 800
251 ThrAlaIleSerGlySerSerLeuAlaHisProGlnGlyLysIleSer 267
801 GACTATGTCACAGTCAGCGCTGTGGCAGGAGCGTGGTGGTACT 850
267 sThrTYrValHisSerAlaValLeuAlaGlyAlaValAlaGlyThr 284
851 CGTGCACCTGATCCCTCTCCGTGGCTGCATGCTGGCTGGCTGG 900
284 eTCyHisIleuLeuLeuProSerProTrpLeuAlaMetValLeuVal 300
901 GCTGGGCTGATCTCCGTGGGAGCCCAAGTACCTGCCGGGTGTGTA 950
301 AlaGlyLeuIleSerValGlyGlyAlaLysTYrLeuProValLysCys 317
951 CCGAGTGTGGGGGATCCCCACAGCTCATGCTGGGCTCAACTCAGCT 1000
317 nArgValLeuGlyIleHisIleSerValMetHisSerIlePheSer 334
1001 TGCTGGTCTGCTTGAGAGATCATCTACATGCTGCTGCTGCTGAT 1050
334 euLeuGlyLeuLeuGlyGlyIleThrTYrIleValLeuLeuValLeu 350
1051 ACCGTGGAGCGCGCAATGCAATGATGGCTTCAGGCTCCCTCAGCAT 1100
351 ThrValTTrpAsnGlyAsnGlyMetIleGlyPheGlnValLeuSer 367
1101 TGGGGAACCTCAGCTGGCCATGCTGATAGCTCTCAGCTGCTGCC 1150
367 eGlyLeuLeuSerLeuAlaIleValIleAlaLeuThrSerGlyLeu 384
1151 CAGGTTTGCTCTTAATCTTAATATGGAAGCACCTCATGAGGCTA 1200
384 hGlyLeuLeuLeuAsnLeuLysIleTrpLysAlaProHisValAla 400
1201 TATTTTGATGACCAAGTTTCTGGAAGTTCTCATTTGGCTGGAT 1250
401 TYrPheAspAspGlnValAlpheTrpLysPheProHisLeuAlaVal 417
1251 T 1251
417 e 417
seq_name: sp_human:Q9UDZ5
seq_documentation_block:
ID Q9UDZ5 PRELIMINARY; PRT; 417 AA.
AC Q9UDZ5;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE RHESUS D CATEGORY VI TYPE III PROTEIN.
GN RHD.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Blood;
RX MEDLINE=9815865; PubMed=9490704;
RA Wagner F.F., Gassner C., Mueller T.H., Schoenitzer D., Schunter F.,
RA Flegel W.A.;
RT "Three molecular structures cause Rhesus D category VI phenotypes with
RT distinct immunohematologic features.";
RL Blood 91:2157-2168(1998).

```

```

RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Blood;
RX MEDLINE=94245182; PubMed=8188244;
RA Cherif-Zahar B., Le Van Kim C., Rouillac C., Raynal V., Cartlon J.,
RA Collin Y.;
RT "Organization of the gene (RHEC) encoding the human blood group RhcBe
RT antigens and characterization of the promoter region.";
RL Genomics 19:68-74(1994).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Blood;
RX MEDLINE=9306356; PubMed=1438298;
RA Le Van Kim C., Mouro I., Cherif-Zahar B., Raynal V., Cherrier C.,
RA Cartlon J.P., Collin Y.;
RT "Molecular cloning and primary structure of the human blood group Rhd
RT polypeptide.";
RL Proc. Natl. Acad. Sci. U.S.A. 89:10925-10929(1992).
DR EMBL; Z97026; CAB09722.1; -
DR InterPro; IPR001905; -
DR InterPro; IPR002229; -
DR Pfam; PF00909; Ammonium transp; 1.
DR PRINTS; PR00342; RHESUSRHD.
SQ SEQUENCE 417 AA; 45247 MW; 1887A3C519149F8F CRC64;

```

```

alignment_scores:
  Quality: 2033.00      Length: 417
  Ratio: 4.959          Gaps: 0
  Percent Similarity: 98.321  Percent Identity: 95.444

```

```

alignment_block:
US-09-600-714-41 x Q9UDZ5

```

```

Align seg 1/1 to: Q9UDZ5 from: 1 to: 417

```

```

1 ATGAGCTTAACTACCCGGCTGTCCGCGCTGCTGCCCTGCTGGGC 50
1 MetSerSerLysTYrProArgSerValArgGlyLeuProLeuTrpAl 17
51 CCTACACCTGGAGAGCAGCTCTCATCTCTCTCTATTTTACCACCT 100
17 AleuThrLeuGlnAlaAlaLeuIleLeuLeuPheTYrPhePheHis 34
101 ATGAGCGCTCTTAGAGATCAAAAGGGGCTGGGCACTATCAAGT 150
34 YrAspAlaSerLeuGlnAspGlnLysGlyLeuValAlaSerTYrGlnVal 50
151 GGCAAGATCTGACCGTGATGGCGGCATTTGGCTTCCACCTC 200
51 GlyGlnAspLeuThrValMetAlaAlaIleGlyLeuGlyPheLeuThr 67
201 GAGTTTCCGAGACACAGCTGAGCAGTGTGGCTTCAACCTTTCATGC 250
67 rSerPheArgArgHisSerTrpSerSerValAlaPheAsnLeuPheMet 84
251 TGGCGCTTGGCTGAGTGGGCAATCCCTGGAGCGGCTCTGGAGCC 300
84 euAlaLeuGlyAlaGlnTrpAlaIleLeuLeuAspLysPheLeuSerGln 100
301 TTCCCTTCTGGAGAGTGTCATCACACTGTTCAGTATTCGGCTGCC 350
101 PheProSerGlyLysValValIleThrLeuPheSerIleArgLeuAla 117
351 CATGAGTGTCTTGTGGCTGCTGATCTCACTGATGGATGCTGCTGG 400
117 rMetSerAlaMetSerValLeuIleSerAlaGlyAlaValLeuGlyLys 134
401 TCAACTGGCGAGTGTGGTGTGATGGTGGCTGGAGGTGACAGTTTA 450
134 AlaAsnLeuAlaGlnLeuValValMetValLeuValGlnValThrAla 150
451 GGCAACCTGAGATGGTGCATCATATATCTTCAACACAGACTACCAT 500

```

```

|||||
151 GYTHrLeuArGMeValIleSerAsnIlePheAsnThrAspTyHisMe 167
501 GAACATGATGACATCTACGTTCGACACCTATTTGGCGTCTGTGCG 550
167 LAsnLeuArGHisPheTyValPheAlaIaIaTyPheGlyLeuThrValA 184
551 CCTGGTGCCTGCCAAAGCCTTACCCGAGGAGGAGGATGAAGATCAG 600
184 IaTrpCysLeuProLysProLeuProLysGlyThrGlnAspAsnSpGln 200
601 ACAGCAGCATACCCAGTTTGTCTGCATGTGGCGCCCTTCTGTG 650
201 ArGAlaThrIleProSerLeuSerAlaMetLeuGlyAlaLeuPheLeuTr 217
651 GATGTTTCGCGCAAGTTTCACTCTGCTCTGTGAGAAATCCAAATCGAA 700
217 PheTrpHeTrpProSerValAsnSerAlaLeuAlaTrpSerProIleGlnA 234
701 GGAAGAAATCCCTGTTCACACCTACTATGCTGTAGACATGACGGTGTG 750
234 rGlyAsnAlaMetPheAsnThrTyTrAlaIaLeuAlaIaSerValVal 250
751 ACAGCCATCTCAGGGTCATCTTGGCTCACCCCAAGGAGATCAGCAA 800
251 ThrAlaIleSerGlySerSerLeuAlaHisProGlnArGlyIleSerMe 267
801 GACTATATGCGACAGTGGCGGTGTGGCAGAGAGCGGTGTGTGGTACT 850
267 tThrTyTrValHisSerAlaValaLeuAlaGlyIaValaIaValaGlyThs 284
851 CGTGTACCTGATCCCTTCTCCGTGGTGCATGCTGTGGGTCTGTG 900
284 eTrCysHisIleLeuIleProSerProTrpLeuAlaMetValLeuVal 300
901 GCTGGCGTGTATCTCCGTGGGAGGAGCAAGTACTCGCGGGGTGTGTA 950
301 AlAGlyLeuIleSerIleGlyIaIaLysCysLeuProGlyCysCysAs 317
951 CCGAGTGGTGGGATCCCCACAGCTCATGTCGGGTACAACTCAGCT 1000
317 nArGValLeuGlyIleProHisSerSerIleMetGlyTrAsnPheserL 334
1001 TGCTGGTCTGCTTGAGAGATCATCTACATTTGCTGCTGCTGCTGAT 1050
334 euLeuGlyLeuLeuGlyGluIleIleTyIleValLeuLeuValLeuAsp 350
1051 ACCGTGCGAGCCGCAATGGCATGTGATGGCTTCCAGGTCTCTCAGCAT 1100
351 ThrValIGlyAlaGlyAsnGlyMetIleGlyPheGlnValLeuLeuSerI 367
1101 TGGGAGACTCAGCTGGCCATGCTGATAGCTCTACAGTCTGTGTCCTGA 1150
367 eGlyIaLeuSerLeuAlaIleValIleAlaLeuThrSerGlyLeuLeuTr 384
1151 CAGGTTTCTCTTAATCTTAAATATGAAAGACACTATGAGGCTAA 1200
384 hGlyLeuLeuLeuAsnLeuLysIleTrpLysAlaProHisGlnAlaLys 400
1201 TATTTTGATGACCAAGTTTCTGGAAGTTTCTCATTTGGCTGTGGAT 1250
401 TyTrPheAspAspGlnValaPheTrpLysPheProHisLeuAlaValaGlyPh 417
1251 T 1251
417 e 417
seq_name: sp_human:Q9UK74
seq_documentation_block:
ID Q9UK74; PRELIMINARY; PRT; 417 AA.
AC Q9UK74;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)

```

```

DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE RH BLOOD GROUP ANTIGEN EVANS.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=PERIPHERAL BLOOD;
RX MEDLINE=96404415; PubMed=8808597;
RA Huang C.H., Chen Y., Reid M., Ghosh S.;
RT "Genetic recombination at the human RH locus: a family study of the
RT red-cell Evans phenotype reveals a transfer of exons 2-6 from the RHD
RT to the RHCE gene."
RL Am. J. Hum. Genet. 59:825-833(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=PERIPHERAL BLOOD;
RA Huang C.H.;
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF17938; AAD5773.1; -.
DR InterPro; IPR001762; -.
DR InterPro; IPR001905; -.
DR InterPro; IPR002229; -.
DR Pfam; PF00909; Ammonium transp; 1.
DR PRINTS; PR00342; RHESUSRD.
DR ProDom; PD000664; -.
SQ SEQUENCE 417 AA; 45445 MW; D8AA0B148ADCT3F7 CRC64;

```

alignment_scores:

Quality:	2016.00	Length:	417
Ratio:	4.990	Gaps:	0
Percent Similarity:	96.882	Percent Identity:	96.403

alignment_block:

US-09-600-714-41 x Q9UK74 ..

Align seg 1/1 to: Q9UK74 from: 1 to: 417

```

1 ATGAGCTCTAAGTACCGCGGTCTGTGTCGGCGCTGCTGCCCTCTGGGC 50
1 MetSerSerLysTyProArGSerValaIaIaGlyCysLeuProLeuCyAl 17
51 CCTACACTGGAAGACGCTCATCTCTCTCTCTATTTTACCCT 100
17 aleuThrLeuGlnAlaAlaLeuIleLeuLeuPheTyTrPheThrHisT 34
101 ATGACGCTTCTTAGAGAGATCAAAAGGGGCTCGTGCAATCTATCAGTT 150
34 yTrAspAlaSerLeuGlnAspGlnLysGlyLeuValaIaSerTyGlnAla 50
151 GAGCAAGATCTGACCGGTGATGGCGCATTTGGCTTGGCTTCTCACCTC 200
51 GlyIaAspLeuThrValaMetAlaIaLeuGlyLeuGlyPheLeuThrSe 67
201 GAGTTTCCGAGACACACAGCTGAGAGCTGAGCGCTTCAACCTTTCATGC 250
67 rSerPheArGArGHisSerTrpSerSerValaIaPheAsnLeuPheMetL 84
251 TGGCGCTTGTGTGACAGTGGCAATCTCTGTGACAGCGCTTCTCAGCCAG 300
84 euAlaLeuGlyValaGlnTrpAlaIleLeuLeuAspGlyPheLeuSerGln 100
301 TTCCTCTTGGGAAGGTGTCATCACACTGTTCAGATTCGGCTGGCCAC 350
101 PheProSerGlyLysValaValaIleThrLeuPheSerIleArGlyLeuAlaTh 117
351 CATGAGTCTTGTGCGTGCATGCTCAGTGCATGCTGCTTGGGGAAG 400
117 rMetSerAlaLeuSerValaLeuIleSerValaIaSpAlaValaLeuGlyLysV 134

```

```

401 TCAACTGGCGCAGTGGTGGTGAATGCTGCTGGAGCTGACAGCTTAA 450
134 aLaenLeuAlaGlnLeuValValMetValLeuValGlnValThrAlaLeu 150
451 GGCAACCTGAGAGATGTCATCATTAATATCTTCAACACAGACTACCAT 500
151 GlysAlaLeuArgMetValIleSerAsnIlePheAsnThrAspTyrHisMe 167
501 GAACATGATGCACATCTACGTGTCGACAGCTATTTTGGCTGCTGGG 550
167 tAsmMetMetHisIleTyrValIlePheAlaIleTyrPheGlyLeuSerVal 184
551 CCGTGGCTGCCAAAGCCTTACCCGAGGAGACGAGATAAAGATCAG 600
184 latrPcysLeuProIysProIysProIysGlnGlyThrGlnAspLysAspIln 200
601 ACAGCAACGATACCCAGTTTGTCTGCCATGCTGGGCGCCCTTCTTG 650
201 ThrIleThrIleProSerLeuSerAlaMetLeuGlyAlaLeuPheLeuTr 217
651 GATGTTCTGGCCAAAGTTTCAACTGCTGCTGCTGAGAGATCCAAATGAA 700
217 pMetPheThrProSerPheAsnSerAlaLeuLeuArgSerProIleGln 234
701 GGAAGATGCCGTGTTCACACCTACTATGCTGTAGACAGTACGCTGTG 750
234 rGlyAsnAlaValIlePheAsnThrTyrTyrAlaValAlaValSerVal 250
751 ACAGCAACGATGAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 800
251 ThrAlaIleSerGlySerSerLeuAlaHisProGlnArgLysIleSerLy 267
801 GACTTATGTGCACAGTGGCTGTTGGCAGAGGCGTGGCTGGTGGTACT 850
267 sThrTyrValHisSerAlaValIleAlaGlyGlyAlaValAlaGlyThr 284
851 CGGTGACCTGATCCCTTCTCCGTGGCTGCCATGCTGCTGGCTGTGTG 900
284 eRcysHisIleuIleProSerProIlePheAlaMetValLeuGlyLeuVal 300
901 GCTGGCTGATCCGTGGGAGGAGCAAGTACCTGCCGGGCTGTGTGAA 950
301 AlaGlyLeuIleSerValGlyGlyAlaLysTyrLeuProValCysCysAs 317
951 CCGAGTGTGGGATTTCCACAGCTCCATGAGGCTACACACTTCACACT 1000
317 nArgValLeuGlyIleHisHisIleSerValMetHisSerIlePheSerL 334
1001 TGCTGGTCTGCTTGGAGATCATCTACATTTGCTGCTGCTGCTGAT 1050
334 euLeuGlyLeuLeuGlyIleThrTyrIleValLeuLeuValLeuHis 350
1051 ACCGTGGAGCCGCAATGCGATGATTTGGCTTCCAGTCTCTCAGCAT 1100
351 ThrValTrpAsnGlyAsnGlyMetIleGlyPheGlnValIleLeuSerI 367
1101 TGGGAACTGAGCTTGGCCATCGTATGCTGCTGCTGCTGCTGCTGCA 1150
367 eGlyGlnLeuSerLeuAlaIleValIleAlaLeuThrSerGlyLeuLeuT 384
1151 CAGGTTTCTCTTAATCTTAATATGAAAGCAGCTCATGAGGCTAA 1200
384 hrglyLeuLeuLeuAsnLeuLysIleTyrLysAlaProHisValAlaLys 400
1201 TATTTTGAATGACCAAGTTTCTGGAAGTTTCTCATTTGGCTGTGAT 1250
401 TyrPheAspAspGlnValIlePheTyrLysPheProHisLeuAlaValGlyPh 417
1251 T 1251
417 e 417
seq_name: sp_human: Q9UPN0

```

```

seq_documentation_block:
ID Q9UPN0 PRELIMINARY; PRT; 417 AA.
AC Q9UPN0;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, last sequence update)
DT 01-MAY-2001 (TREMBlrel. 16, last annotation update)
DE RH BLOOD CE GROUP ANTIGEN POLYPEPTIDE.
GN RHE.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Kashlase K., Ishikawa Y., Uchikawa M.;
RT "The serological profile and molecular basis of the RHE variant in
RT Japanese.";
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Kashlase K., Uchikawa M., Ishikawa Y.;
RT "Rh blood CE group antigen polypeptide, RHEK.";
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB030388; BA82627.1; -.
DR EMBL; AB049753; BAB16597.1; -.
DR InterPro; IPR001905; -.
DR InterPro; IPR002229; -.
DR Pfam; PF00909; Ammonium transp; 1.
DR PRINTS; PR00342; RHESUSRD.
SO SEQUENCE 417 AA; 45604 MW; 61E4BB39ABA23CC CRC64;

```

```

alignment_scores:
Quality: 1963.00 Length: 417
Ratio: 4.920 Gaps: 0
Percent Similarity: 95.683 Percent Identity: 93.285

```

```

alignment_block:
US-09-600-714-41 x Q9UPN0

```

```

Align seg 1/1 to: Q9UPN0 from: 1 to: 417

```

```

1 ATGAGCTCTAAGTACCCGCGTCTGTCGCGGCTGCTGCCCTCTGGGC 50
1 MetSerSerLysTyrProArgSerValArgArgCysLeuProLeuTrpAl 17
51 CCTAACCTGGAAGCAGCTTCATTCCTCTCTTATTTTATCCCACT 100
17 AleuThrLeuGlnAlaAlaLeuIleLeuLeuPheTyrPhePheThrHisT 34
101 ATGAGCTCTCTTACAGATCAAAAGGGGCTGCGCATCCATCAATGTT 150
34 TyrAspAlaSerLeuGlnAspGlnLysGlyLeuValAlaSerTyrGlnVal 50
151 GGGCAAGATCTGACCGTATGAGCGGCATTTGGCTTGGCTTCCACCTC 200
51 GlyGlnAspLeuThrValMetAlaAlaIleGlyLeuGlyPheLeuThrSe 67
201 GAGTTTCGGAGACACAGCTGAGAGAGTGTGGCTTCAACCTCTTCATGC 250
67 rSerPheArgArgHisSerTyrSerSerValAlaIlePheAsnLeuPheMetL 84
251 TGGCGCTGTGTGTCAGTGGGCAATCCCTGCTGAGAGGCTCTCTAGCCAG 300
84 euAlaLeuGlyValGlnTrpAlaIleLeuLeuAspGlyPheLeuSerGln 100
301 TTCCCTCTGGAAGGTGTCATCACACTGTTCATATTCGCTGGCTGGCCAC 350
101 PheProSerGlyLysValValIleThrIlePheSerIleIleArgLeuAlaTh 117
351 CATGAGTCTTTTGGTGGTGCATCTCAGTGGATGCTCTTGGGGAAG 400

```

```

117 mterSerAlaLeuSerValLeuIleSerValAspAlaValLeuGlyLysV 134
401 TCACTTGGCCGCACTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 450
134 aIaenLeuAlaGlnLeuValValMetValLeuValGlnValThrAlaLeu 150
451 GGCAACCGGAGGATGTCATCAATATCTTCAACACAGACATCAACAT 500
151 GlnsLeuArgMetValIleSerAsnIlePheAsnThrAspTyrHisMe 167
501 GAACATGATGCACATCTACGTGGCAGCCCTAATTTGGGCTGTCTGG 550
167 tAsnLeuArgHisPheTyrValPheAlaAlaTyrPheGlyLeuThrValA 184
551 CCTGTGGCTGCCAAGCCCTACCCGGAAGGAGGAGGAGGAGGAGGAGGAG 600
184 larrPcysLeuProLysProLeuProLysGlyThrGlnAspAsnAspGln 200
601 ACAGCAACGATACCCAGTTGTGTGCTGCCATGCTGGGCGCCCTTCTGTG 650
201 ArgAlaThrIleProSerLeuSerAlaMetLeuGlyAlaLeuPheLeuTr 217
651 GATGTTCTGCCAAGTTCAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 700
217 pMetPheTrpProSerValAsnSerProLeuLeuArgSerProIleGlnA 234
701 GGAGGATGCCGCTGTCAACACCTACTATGCTGTAGCAGTACAGTGGTG 750
234 rGlyAsnAlaMetPheAsnThrTyrTyrAlaLeuAlaValSerValVal 250
751 ACAGCAACCTCAGGATCATCCTTGGCTACCCGCAAGGAGGAGGAGGAGG 800
251 ThrAlaIleSerGlySerSerLeuAlaHisProGlnArgLysIleSerMe 267
801 GACTTATGTGCACAGTGGCTGGTGGCAGGAGGCTGGCTGGCTGGTACT 850
267 tThrTyrValHisSerAlaValLeuAlaGlyValAlaValAlaGlyThrS 284
851 CGTGCACCTGATCCCTTCCGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 900
284 eRcysHisLeuIleProSerProTrpLeuAlaMetValLeuGlyLeuVal 300
901 GCTGGGCTGATCTCCGCGGGGAGCCAACTACCTCCGCGGGCTGTGTA 950
301 AlaGlyLeuIleSerIleGlyAlaLysCysLeuProValCysCysAs 317
951 CCGAGTGTGGGATTCGCCACAGCTCATGATGGCTGACAACTTCAGCT 1000
317 nArgValLeuGlyIleHisHisIleSerValMetHisSerIlePheSerL 334
1001 TGCTGGGCTCTGGAGAGATCATCTACATGTGCTGGCTGGCTGGTGGT 1050
334 euLeuGlyLeuLeuGlyGlnIleThrTyrIleValLeuLeuValLeuHis 350
1051 ACCGTCGAGAGCCGCAATGGCATGATGGCTTCAGGTCCTCCTCAGAT 1100
351 ThrValTrpAsnGlyAsnGlyMetIleGlyPheGlnValIleLeuSerI 367
1101 TGGGGAACTACGCTTGGCCATCGATAGCTCTCAGCTGTGCTCTCGGA 1150
367 eGlyGlnLeuSerLeuAlaIleValIleAlaLeuThrSerGlyLeuLeu 384
1151 CAGGTTTGCTCTTAATCTTAATATGGAAGACCTCATGAGGCTTAA 1200
384 hrGlyLeuLeuLeuAsnLeuLysIleTrpLysAlaProHisValAlaLys 400
1201 TATTTTGATGACCAAGTTTCTGGAAGTTTCTCATTTGGCTGTGGATT 1250
401 TyrPheAspAspArgIleValPheTrpLysPheProHisLeuAlaValGlyPh 417
1251 T 1251
417 e 417

```

```

seq_name: sp_human:Q9UEC3
seq_documentation_block:
ID Q9UEC3 PRELIMINARY; PRT; 417 AA.
AC Q9UEC3;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE RH BLOOD CE GROUP ANTIGEN POLYPEPTIDE.
GN RHCE.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homindae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Kashwase K., Ishikawa Y., Uchikawa M.;
RT "The serological profile and molecular basis of the RHE variants
RT (RHEp and RHEKH) in Japanese.";
RL Submitted (OCT-1998) to the EMBL/Genbank/DBJ databases.
DR EMBL; AB018644; BAA33927.1; -.
DR InterPro; IPR000885; -.
DR InterPro; IPR001905; -.
DR InterPro; IPR002229; -.
DR Pfam; PF00909; Ammonium transp. 1.
DR PRINTS; PR00342; RHESUSRHD.
DR PRODOM; PD002078; -.
DR SEQUENCE 417 AA; 45529 MW; 3B072CA1544F48C0 CRC64;

alignment_scores:
Quality: 1940.00 Length: 417
Ratio: 4.887 Gaps: 0
Percent Similarity: 95.204 Percent Identity: 92.086

alignment_block:
US-09-600-714-41 x Q9UEC3 ..
Align seg 1/1 to: Q9UEC3 from: 1 to: 417

1 ATGAGCTTAAGTACCCGGGCTGTCCGGGCTGCCGCCCTTGCGC 50
1 MetSerLysTyrProArgSerValArgArgCysLeuProLeuTrpAl 17
51 CCTACACTGGAGCAGCTCATCTCCCTCTATTTTACCACCT 100
17 aLeuThrLeuGlnAlaIleLeuIleLeuLeuPheTyrPheThrHisT 34
101 ATGACGCTTCCTTAGAGATCAAAAGGGGCTCGTGCATCCTATCAAGTT 150
34 yAspAlaSerLeuGlnAspGlnLysGlyLeuValAlaSerTyrGlnVal 50
151 GGCCAAGATTCGACCGTGTATGGCGGCATTTGGCTTCTCACCTC 200
51 GlnGlnAspLeuThrValMetAlaAlaLeuGlyLeuGlyPheLeuTrpS 67
201 GAGTTCCGGAGACACAGCTGGAGCAGTGGGCTTCAACCTTTCATGC 250
67 tAsnPheArgArgHisSerTrpSerSerValAlaPheAsnLeuPheMetL 84
251 TGGCGTTGGTGTGTCAGTGGGCAATCCTGTGAGCGGCTTCGAGGCAG 300
84 euAlaLeuGlyValGlnTrpAlaIleLeuLeuAspGlyPheLeuSerGln 100
301 TTCCTTCTGGGAAGGTGTCATCACACTGTTCAGTATTTGGCTGGCCAC 350
101 PheProProGlyLysValValIleThrLeuPheSerIleArgLeuAlaThr 117
351 CATGAGGCTTGTGCGTGGCTGATCGATCGATGAGATCGTGTGGGAGG 400
117 mterSerAlaLeuSerValLeuIleSerValAlaValLeuGlyLysV 134

```

seq_name: sp_human:Q9UEC2

```

401 TCAACTGGCGCACTGGTGGTGCATGCTGCTGGTGGAGAGTCAAGCTTTA 450
134 aAlaLeuAlaGlnLeuValValMetValLeuValGluValThrAlaLeu 150
451 GGCACACCTGAGAGTGCATGATCAATATATCTTCAACACAGCTCCACAT 500
151 G1YThrLeuArgMetValIleSerAsnIlePheAsnThrAspTyrHisSme 167
501 GAACATGATGCACATCTACGTGTTCGACAGCCTATTTGGCTGCTGTGG 550
167 tAsnLeuArgHisPheTyrValPheAlaIaLalTyrPheGlyLeuThrValA 184
551 CCTGGCTGCTGCCAAAGCCTTACCCGAGGAGACGAGATAAAGATCAG 600
184 latrPcysLeuProlyspProleuProlysglyThrGluAspAsnspcIn 200
601 ACAGCAAGATACCCAGTTGCTGTCGCCATGCTGGCGCCCTCTCTGTG 650
201 ArgAlaThrIleProSerLeuSerAlaMetLeuGlyAlaLeuPheLeuTr 217
651 GATGTTTGCCCAAGTTTCAACTGCTGCTGCTGAGAACTCAATCGAAA 700
217 pMetPheTrpProSerValAsnSerProleuLeuArgSerProIleGluA 234
701 GGAAGAAATGCCGTGTTCACACCTACTAGTGTGATGACAGTCAGCGTGG 750
234 rGlyAsnAlaValPheAsnThrTyrTyrAlaLeuAlaValSerValVal 250
751 ACAGCCATCTCAGGGTGCATCCTGCTGCTACCCCAAGGAGATCAGCAA 800
251 ThrAlaIleSerGlySerSerLeuAlaHisProGlnArgLysIleSerMe 267
801 GACTTATGTGCACAGTGGGTGTGGTGGCAGAGCGGTGCTGGGTACT 850
267 tThrTyrValHisSerAlaValLeuAlaGlyGlyAlaValAlaGlyThrS 284
851 CGTGTACCTGATCCCTTCCGTGGCTGGCATGCTGCTGGCTGTGTG 900
284 eRcYHisIleuIleProSerProIleuAlaMetValLeuGlyLeuVal 300
901 GCTGGGCTGATCTCCGTGGGGGAGCCAAAGTACCTGCCGGGTGTGTGA 950
301 AlaGlyLeuIleSerIleGlyGlyAlaLysCysLeuProValLysCysAs 317
951 CCGAGTGTGGGATGCCACAGCTCCATCATGGGCTACAACTTCACACT 1000
317 nArgValLeuGlyIleHisIleSerValMetHisSerIlePheSerL 334
1001 TGCTGGTCTGCTGGAGAGATCATCTACATTTGCTGCTGGTCTGTAT 1050
334 euLeuGlyLeuLeuGlyGluIleThrTyrIleValLeuLeuValLeuHis 350
1051 ACCGTGGAGCGCGCAATGCGATGATTTGGCTTCCAGGTCTCTCAGAT 1100
351 ThrValIlePheAsnGlyAsnGlyMetIleGlyPheGlnValLeuLeuSerI 367
1101 TGGGGAATCAGCTTGGCCATGATGATAGCTCTCAGTCTGCTGCTCTGA 1150
367 eGlyIuLeuSerLeuAlaIleValIleAlaLeuThrSerGlyLeuLeuT 384
1151 CAGGTTGCTCTTAATCTTAATAATATGAAAGACACTCATGAGGCTAAA 1200
384 nrgIyLeuLeuLeuAsnLeuLysIleTrpLysAlaProHisValAlaLys 400
1201 TATTTTGATGACAGTTTCTCGAAGTTTCTCATTTGGCTGGCTGGAT 1250
401 TyrPheAspAspGlnValPheTrpLysPheProHisLeuAlaValGlyPh 417
1251 T 1251
417 e 417

```

```

seq_documentation_block:
ID Q9UEC2 PRELIMINARY; PRT; 417 AA.
AC Q9UEC2;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE RH BLOOD GROUP CE ANTIGEN POLYPEPTIDE.
GN RHCE.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Kashiwase K., Ishikawa Y., Uchikawa M.;
RT "The serological profile and molecular basis of the RHE variants
RT (RHEFM and RHEKH) in Japanese.";
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB018645; BAA33928.1; -.
DR InterPro; IPR001905; -.
DR InterPro; IPR002229; -.
DR Pfam; PF00909; Ammonium transp; 1.
DR PRINTS; PR00342; RHESUSRD.
SQ SEQUENCE 417 AA; 45505 MW; 67B53E74CB053DE CRC64;

```

alignment_scores:

```

Quality: 1928.00 Length: 417
Ratio: 4.869 Gaps: 0
Percent Similarity: 94.964 Percent Identity: 91.367

```

alignment_block:

```

US-09-600-714-41 x Q9UEC2 ..

```

```

Align seg 1/1 to: Q9UEC2 from: 1 to: 417

```

```

1 ATGAGCTTAAGTACCCGCGGTGTGCTCCGCGCTGCTGCCCTCTGGGC 50
1 MetSerLeuLysTyrProArgSerValArgArgCysLeuProLeuTrpAl 17
51 CCTAAGCTGGAAGCAGCTTCATTTCTCTCTCTATTTTATTTTACCCACT 100
17 aLeuThrLeuGlnAlaIleAlaLeuIleLeuLeuPheTyrPhePheThrHis 34
101 ATGAGCTTCTCTAGAGATCAAAAGGGCTCGTGGCTCATCTCAAGATT 150
34 yAspAlaSerLeuGlnAspGlnLysGlyLeuValAlaSerTyrGlnVal 50
151 GGCACAAGATCGACCGTATGCGCGCATTTGACTTGGGCTTCTCACCTC 200
51 GlyGlnAspLeuThrValMetAlaIleAlaLeuGlyLeuGlyPheLeuThrSe 67
201 CAGTTTCCGGAGACACAGCTGAGCAGTGTGGCTTCAACCTCTTCAGCC 250
67 rAsnPheArgArgHisSerTrpSerSerValAlaIlePheAsnLeuPheMet 84
251 TGGGCGTTGGTGTGCAAGTGGGCAATCTGCTGGAGCGGCTTCTGAGCC 300
84 euAlaLeuGlyValGlnTrpAlaIleLeuLeuAspGlyPheLeuSerGln 100
101 PheProPheGlyLysValValIleThrLeuPheSerIleArgLeuAlaTh 117
301 TTCCCTTGTGGGAAGTGGTCAATCAGACTGTTCACTATTCGCTGGCCAC 350
351 CATGAGTCTTTGTGCTGCTGCATCTCAGATGATGCTGTCTTGGGGAAG 400
117 lMetSerAlaMetSerValLeuIleSerAlaGlyAlaValLeuGlyLysV 134
401 TCAACTGGCGCACTGGTGGTGCATGCTGCTGCTGAGAGTCAAGCTTTA 450
134 aAlaLeuAlaGlnLeuValValMetValLeuValGluValThrAlaLeu 150

```

```
451 GGCACCTGAGGATGTCATCATTAATATCTTCAACACAGACTACACAT 500
|||||
151 GlyThrLeuThrMetValIleSerAsnIlePheAsnThrAspTyrHisMe 167
501 GAACATGATGCACATCTACGTGTTCCGACGCTATTGGCTGTGTGG 550
|||||
167 LAsnLeuArgHisPheTyrValPheAlaIaIaTyrPheGlyLeuThrValA 184
551 CCTGTGCTCCGCAAGGCTCTACCCGGGAGGAGGAGGATAAAGATCAG 600
|||||
184 LArgPcysLeuProLysProLeuProLysGlyThrGluAspAsnSpGln 200
601 ACAGCAACGATACCCAGTTGTCTGCCATGCTGGCGGCCCTCTTGTG 650
|||||
201 ArgAlaThrIleProSerLeuSerAlaMetLeuGlyAlaLeuPheLeuTr 217
651 GATGTTCTGGCCAAAGTTTCAACTGCTGCTGTGAGAAAGTCCAAATCGAA 700
|||||
217 pMetPheTrpProSerValAsnSerProLeuLeuArgSerProIleGlnA 234
701 GGAAGATGCCGTGTCAACACCTACTAGTGTGAGCAGTACGCGGTGG 750
|||||
234 rGlySAsnAlaMetPheAsnThrTyrTrpAlaLeuAlaValSerValVal 250
751 ACAGCCATCTCAGGGTCATCCTGGCTCACCCCAAGGAGAGATCAGCAA 800
|||||
251 ThrAlaIleSerGlySerSerLeuAlaHisProGlnArgLysIleSerMe 267
801 GACTTATGTGCACAGTCGCGGTGTGGCAGAGGCGGTGGCTGGGTACT 850
|||||
267 tThrTyrValHisSerAlaValLeuAlaGlyGlyValAlaValAlaGlyThrS 284
851 CGTGTACACTGATCCCTTCCGTGGCTTGGCATGATGCTGGTCTGTG 900
|||||
284 e rCysHisLeuIleProSerProTrpLeuAlaMetValLeuGlyLeuVal 300
901 GCTGGGCTGATCTCCGTGGGGGAGCAAGTACTCCCGGGGTGTGTA 950
|||||
301 AlaGlyLeuIleSerIleGlyGlyAlaLysCysLeuProValCysCysAs 317
951 CCGAGTGTGGGGATCCCCACAGCTCATCATGGCTACACTTCAGCT 1000
|||||
317 nArgValLeuGlyIleHisHisIleSerValMetHisSerIlePheSerL 334
1001 TGCTGGGCTGCTGGAGAGATCATCTACATTTGCTGCTGCTGTGAT 1050
|||||
334 e uLeuGlyLeuLeuGlyGlyIuIleThrTyrIleValLeuLeuValLeuHis 350
1051 ACCGTGAGGACCGGCAATGGCATGATTGGCTTCCAGGTCTCTCAGCAT 1100
|||||
351 ThrValITrpAsnGlyAsnGlyMetIleGlyPheGlnValLeuLeuSerI 367
1101 TGGGGAACCTACGCTTGGCCATCGTATAGCTCTACAGTCTGCTCCGTA 1150
|||||
367 eGlyGluLeuSerLeuAlaIleValIleAlaLeuThrSerGlyLeuLeuT 384
1151 CAGGTTTGCTCTTAATCTTAAATATGGAAGACACCTCATGAGGCTAAA 1200
|||||
384 h rGlyLeuLeuLeuAsnLeuLysIleTrpLysAlaProHisValAlaLys 400
1201 TATTTTGATGACCAAGTTTCTGGAAGTTTCCATTTGGCTGTGGATT 1250
|||||
401 TyrPheAspAspGlnValPheTrpLysPheProHisLeuAlaValGlyPh 417
1251 T 1251
417 e 417
```

OM of: US-09-600-714-41 to: SwissProt_39:* out_format : pfs

Date: Sep 11, 2001 8:53 AM

About: Results were produced by the Gencore software, version 4.5,
Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:

```
-MODEL=frame_n2p.model -DEV=xlp  
-O=cgcn2.1/USPRO.spool/US09600714/runat_11092001_083825_22207/app-query.fasta.1.1332  
-DB=SwissProt_39 -OFMT=fasta -SUFFIX=isp -GAPOP=12.000  
-GAPEXT=4.000 -MINMATCH=0.100 -LOOPEXT=0.000 -LOOPEXT=0.000  
-GAPOP=4.500 -GAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500  
-FGAPOP=6.000 -FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500  
-DELDP=6.000 -DELDP=7.000 -START=1 -MATRIX=bloms62  
-TRAN=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pc  
-THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs  
-NORMEXT=0 -MIXLEN=200000000  
-USER=US09600714_ECON1_1_34 -NCPU=6 -ICPU=3 -LONGLOG -NO_XLPRX  
-WAIT -THREDS=1
```

Search information block:

```
Query: US-09-600-714-41  
Query length: 1254  
Database: SwissProt_39:*  
Database sequences: 93435  
Database length: 34255486  
Search time (sec): 30.140000
```

score_list:

```
Sequence      Strd Orig      ZScore      EScore Len      Documentation  
SwissProt_39:RHD_HUMAN + 2118.00 2703.61 1.4e-143 416 020161 homo sapiens (human).  
SwissProt_39:RHLA_PATNR + 1970.00 2514.36 5.0e-133 416 028814 pan troglodytes (chm  
SwissProt_39:RHLA_GONGO + 1948.00 2486.23 1.8e-131 416 028814 pan troglodytes (chm  
SwissProt_39:RHLA_PATNR + 1947.00 2484.95 2.2e-131 416 028813 pan troglodytes (chm  
SwissProt_39:RHLA_HUMAN + 1929.00 2461.93 4.2e-130 416 018577 homo sapiens (human).  
SwissProt_39:RHLF_PATNR + 1914.00 2442.75 4.9e-129 416 028812 pan troglodytes (chm  
SwissProt_39:RHLG_GONGO + 1903.00 2428.69 3.0e-128 416 028842 gorilla gorilla goril  
SwissProt_39:RHLA_MACMU + 1642.00 2094.94 1.2e-109 416 028849 macaca mulatta (rhesu  
SwissProt_39:RHLA_MACRA + 1634.00 2084.71 4.3e-109 416 028841 macaca fascicularis (c  
SwissProt_39:RHLA_HYLT + 1527.00 1949.30 1.8e-101 353 020446 hylobates pileatus (pi  
SwissProt_39:RHLA_HUMAN + 628.50 799.12 1.8e-37 409 020446 homo sapiens (human).  
SwissProt_39:SRCH_HUMAN + 141.00 171.16 0.0098 699 023327 homo sapiens (human).  
SwissProt_39:Y108_SYNY3 + 138.50 170.72 0.0143 507 023327 synchocystis sp. (st  
SwissProt_39:RHUB_ECOLI + 136.50 165.90 0.0204 660 060972 escherichia coli. f  
SwissProt_39:MOKD_SCHPO + 135.50 158.71 0.0272 238 039719 schizosaccharomyces  
SwissProt_39:COX1_APLI + 133.50 163.45 0.0354 521 020374 apis mellifera ligust  
SwissProt_39:MY15_MOUSE + 130.50 143.90 0.0644 351 029424 mus musculus (mouse)  
SwissProt_39:SRCH_RABIT + 129.00 154.12 0.0715 852 016230 onychomys leucogaster  
SwissProt_39:AMT1_HUMAN + 128.00 157.39 0.0800 501 054144 arabidopsis thaliana  
SwissProt_39:REPT_MOUSE + 123.50 137.16 0.1977 2715 090066 homo sapiens (human)  
SwissProt_39:REPT_MOUSE + 123.00 144.03 0.1968 1130 097347 mus musculus (mouse)  
SwissProt_39:ARB2_ECOLI + 120.50 149.13 0.2695 429 052146 escherichia coli. ars  
SwissProt_39:OCCT6_MOUSE + 120.00 148.10 0.2938 449 021952 mus musculus (mouse).  
SwissProt_39:OCCT6_RAT + 120.00 148.06 0.2940 451 020267 rattus norvegicus (rat)  
SwissProt_39:AMTB_ECOLI + 117.50 145.31 0.4406 428 037905 escherichia coli. p  
SwissProt_39:NUOM_RICPR + 116.50 142.86 0.5283 491 092900 rickettsia prowazeki  
SwissProt_39:AMT_MYCTU + 116.00 142.47 0.5696 477 010968 mycobacterium tuberculosis  
SwissProt_39:IE18_PRYKA + 115.50 132.33 0.6900 1446 033479 pseudotuberculis virus (c  
SwissProt_39:MEP2_YEAST + 115.00 140.80 0.6741 499 041948 saccharomyces cerevis  
SwissProt_39:INVO_RAT + 115.00 139.69 0.6828 568 048998 rattus norvegicus (rat)  
SwissProt_39:HYFB_ECOLI + 115.00 138.25 0.6943 672 034382 escherichia coli. hyc  
SwissProt_39:MY15_HUMAN + 115.00 124.04 0.8181 3530 090407 homo sapiens (human)  
SwissProt_39:YDHC_ECOLI + 114.50 141.99 0.7164 403 037937 escherichia coli. hyl  
SwissProt_39:YAI1_SYNY3 + 114.00 141.20 0.7230 442 017295 synchocystis sp. (st  
SwissProt_39:CHRA_ALCEU + 114.00 141.40 0.7773 401 017551 alcaligenes eutrophus  
SwissProt_39:OCCT6_HUMAN + 113.50 139.81 0.8530 448 003052 homo sapiens (human).  
SwissProt_39:NDAM_MERSE + 113.50 139.02 0.8608 491 047497 metridium senile (brc  
SwissProt_39:MEP1_YEAST + 113.50 139.00 0.8609 492 040260 saccharomyces cerevis  
SwissProt_39:Y092_HAIBN + 113.00 139.74 0.9198 419 057433 haemophilus influenza  
SwissProt_39:GNTP_ZYMMO + 113.00 139.11 0.9265 451 092670 zymomonas mobilis. g
```

```
SwissProt_39:ARSB_YEREN + 112.50 138.90 1.00 429 074985 yersinia enterocol  
SwissProt_39:YDEE_ECOLI + 112.00 138.97 1.08 395 031126 escherichia coli.  
SwissProt_39:GNTT_ECOLI + 111.50 137.46 1.18 437 033935 escherichia coli.  
SwissProt_39:YCDG_ECOLI + 111.50 137.37 1.18 442 075892 escherichia coli.  
SwissProt_39:HRG_HUMAN + 111.50 135.89 1.20 525 040196 homo sapiens (human
```

seq_name: SwissProt_39:RHD_HUMAN

```
seq_documentation_block:  
ID RHD_HUMAN STANDARD: PRT; 416 AA.  
AC 002161: 002162: 007618: 016355: 016147: 016235:  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 01-OCT-2000 (Rel. 40, Last annotation update)  
DE BLOOD GROUP RH(D) POLYPEPTIDE (RHEUSUS D ANTIGEN) (RH  
DE POLYPEPTIDE 2) (RHDII).  
GN RHD.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
OX NCBI_TaxID=9606;  
RN 11  
RP SEQUENCE FROM N.A.  
RC TISSUE=Bone marrow;  
RX MEDLINE=93066356; PubMed=1438298;  
RA le van Kim C., Cherif-Zahar B., Raynal V., Cherrier C.,  
RA Cartron J.-P., Colin Y.;  
RT "Molecular cloning and primary structure of the human blood group Rhd  
RT polypeptide."  
RL Proc. Natl. Acad. Sci. U.S.A. 89:10925-10929(1992).  
RN 12  
RP SEQUENCE FROM N.A.  
RC TISSUE=Bone marrow;  
RX MEDLINE=93260855; PubMed=1379850;  
RA le van Kim C., Cherif-Zahar B., Raynal V., Mouro I., Lopez M.,  
RA Cartron J.-P., Colin Y.;  
RT "Multiple Rh messenger RNA isoforms are produced by alternative  
RT splicing."  
RL Blood 80:1074-1078(1992).  
RN 13  
RP SEQUENCE FROM N.A.  
RX MEDLINE=93320449; PubMed=8329718;  
RA Arce A.A., Thompson E.S., Wagner S., Coyne K.E., Ferdman B.A.,  
RA Lublin D.M.;  
RT "Molecular cloning of a Rhd cDNA derived from a gene present in Rhd-  
RT positive, but not Rhd-negative individuals."  
RL Blood 82:651-655(1993).  
RN 14  
RP SEQUENCE FROM N.A.  
RX MEDLINE=93216282; PubMed=7916743;  
RA Kaji E., Umenishi F., Iwamoto S., Ikemoto S.;  
RT "Isolation of a new cDNA clone encoding an Rh polypeptide associated  
RT with the Rh blood group system."  
RL Hum. Gene. 91:157-162(1993).  
RN 15  
RP SEQUENCE FROM N.A.  
RX MEDLINE=95329738; PubMed=7606008;  
RA Huang C.H., Reid M.E., Chen Y.;  
RT "Identification of a partial internal deletion in the Rh locus  
RT causing the human erythrocyte D-phenotype."  
RL Blood 86:784-790(1995).  
RN 16  
RP SEQUENCE FROM N.A. (SHORT FORM 1).  
RX MEDLINE=94235883; PubMed=8180407;  
RT "Identification of a new Rhd-specific mRNA from K562 cells."  
RL Blood 83:3098-3100(1994).  
RN 17  
RP SEQUENCE FROM N.A. (SHORT FORM 2).  
RX MEDLINE=94362249; PubMed=8080999;  
RT Suyama K., Lunn R., Haller S., Goldstein J.;  
RT "Rh(D) antigen expression and isolation of a new Rh(D) cDNA isoform  
RT in human erythroleukemic K562 cells.";
```

RL Blood 84:1975-1981(1994).

RN (8)

RP VARIANT BLOOD GROUP TAR.

RX MEDLINE-95259709; PubMed-7741145;

RA Rouillac C., Le van Kim C., Beolet M., Carton J.-P., Colin Y.,

RT "Leu110p substitution in the Rhd polypeptide is responsible for the

RT DVI1 category blood group phenotype.";

RL Am. J. Hemtol. 49:87-88(1995).

CC -1- FUNCTION: MAY BE PART OF AN OLIGOMERIC COMPLEX WHICH IS LIKELY TO

CC HAVE A TRANSPORT OR CHANNEL FUNCTION IN THE ERYTHROCYTE MEMBRANE.

CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.

CC -1- ALTERNATIVE PRODUCTS: VARIOUS FORMS ARE PRODUCED BY ALTERNATIVE

CC SPLICING.

CC -1- TISSUE SPECIFICITY: RESTRICTED TO TISSUES OR CELL LINES EXPRESSING

CC ERYTHROID CHARACTERS.

CC -1- POLYMORPHISM: RHD AND RHCE ARE RESPONSIBLE FOR THE RH BLOOD GROUP

CC SYSTEM. THE MOLECULAR BASIS OF THE TAR-RH40 BLOOD GROUP ANTIGEN

CC IS A POLYMORPHISM IN POSITION 109.

CC -1- SIMILARITY: BELONGS TO THE RH FAMILY. STRONG, TO RHCE.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@sib-sib.ch).

CC -----

DR EMBL: X63097; CAA44831.1; -

DR EMBL: X63094; CAA44808.1; -

DR EMBL: L08429; AAA02679.1; -

DR EMBL: S57971; AAB26081.1; -

DR EMBL: S78509; AAB34852.1; -

DR EMBL: S70174; AAB30756.1; -

DR EMBL: S73913; AAB31911.1; -

DR PIR: S26564; S26564.

DR MIM: 111680; -

DR InterPro: IPR001905; -

DR Pfam: PF00909; Ammonium transp; 1.

DR PRINTS: PR00342; RHESUSRHD.

KM Erythrocyte; Transmembrane; Blood group antigen; Alternative splicing;

KW Polymorphism.

FT INIT_MET 0

FT TRANSMEM 11 31 POTENTIAL.

FT TRANSMEM 43 63 POTENTIAL.

FT TRANSMEM 75 96 POTENTIAL.

FT TRANSMEM 106 126 POTENTIAL.

FT TRANSMEM 129 149 POTENTIAL.

FT TRANSMEM 166 186 POTENTIAL.

FT TRANSMEM 202 222 POTENTIAL.

FT TRANSMEM 237 257 POTENTIAL.

FT TRANSMEM 286 306 POTENTIAL.

FT TRANSMEM 333 353 POTENTIAL.

FT TRANSMEM 357 377 POTENTIAL.

FT VARSPPLIC 313 408 MISSING (IN SHORT ISOFORM 1).

FT VARSPPLIC 315 315 C -> S (IN SHORT ISOFORM 2).

FT VARSPPLIC 316 416 MISSING (IN SHORT ISOFORM 2).

FT VARIANT 109 109 L -> P (IN TAR ANTIGEN).

FT VARIANT 109 109 /FTID-VAR_006919.

FT VARIANT 217 217 M -> I.

FT VARIANT 217 217 /FTID-VAR_006920.

FT CONFLICT 15 15 W -> C (IN REF. 5).

FT CONFLICT 38 38 E -> G (IN REF. 4).

FT CONFLICT 102 102 S -> P (IN REF. 4).

FT CONFLICT 126 126 V -> A (IN REF. 4).

FT CONFLICT 173 173 V -> M (IN REF. 5).

FT CONFLICT 181 181 S -> T (IN REF. 4).

FT CONFLICT 313 313 G -> V (IN REF. 4 AND 7).

FT CONFLICT 322 322 P -> H (IN REF. 4).

FT CONFLICT 397 397 E -> V (IN REF. 5).

SO SEQUENCE 416 AA: 45049 MW: 9136DFIA37D76B1B CRC64;

alignment_scores:

Quality: 2118.00 Length: 416

Ratio: 5.091 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

us-09-600-714-41 x RHD_HUMAN ..

Align seg 1/1 to: RHD_HUMAN from: 1 to: 416

4 AGCTGTAGTACCGCGGCTGTCGGCGCTGCTGCCCTTGCGGCTT 53

1 SerSerIstYrProAArgSerValAArgCysLeuProLeuTrpAlaLe 17

54 AACACGTGGAAGCAGCTCTCATCTCCCTCTTCTATTTTAAACCATG 103

17 uThLeuGlnAlaAlaLeuLeuLeuLeuPheTyPhePheThrIstYrA 34

104 ACGCTTCCTTAGAGATCAAAAGGGCTCGTGCATCTCTATCAAGTGGC 153

34 spAlaSerLeuGlnAspGlnIstYleuValAlaSerTyrglnValGly 50

154 CAAGATCTGACCGTGATGGCGCCATGGCTTGGGCTTCTCAGCTGAG 203

51 GlnAspLeuThrValMetAlaAlaIleGlyLeuGlyPheLeuThrSerS 67

204 TTTCCGGAGACACAGCTGGAGCAGTGGCGCTTGCATCTCTCATGCTGG 253

67 rPheArgArgHisSerTrpSerSerValAlaPheAsnLeuPheMetLeuA 84

254 CGCTTGTGTGTCAGTGGGCAATCCTGTGGACGGCTTCTGAGCCAGTTC 303

84 IalLeuGlyValGlnTrpAlaIleLeuLeuAspGlyPheLeuSerGlnPhe 100

304 CCTTCTGGGAAGGTGTCATCACACTTTCAGTATTTGGGCTGGCCACAT 353

101 ProSerIstYsValValIleThrLeuPheSerIleArgLeuAlaTrpMe 117

354 GAGTCTTTTGTGCGTGTGATCTCAGTGGATGAGTGTGTGGGCAAGGTCA 403

117 tSerAlaLeuSerValLeuIleSerValAspAlaValLeuGlyIstYsValA 134

404 ACTTGGCGCATGTTGGTGTATGGTGTGTGTGAGGTGACAGCTTTAGGC 453

134 snLeuAlaGlnLeuValValMetValLeuValGluValThrAlaLeuGly 150

454 AACCTGAGATGTCATCAGTAATATCTTCAACACAGACTACCAATGAA 503

151 AsnLeuArgMetValIleSerAsnIlePheAsnThrAspTyrlHisMetAs 167

504 CATGATGCACATCTACGTTGTCGAGCCTATTTGGGCTGCTGTGGCCT 553

167 nMetMetHisIleTyValPheAlaAlaIstYrPheGlyLeuSerValAlat 184

554 GGTGCTGCCAAAGCCTCTACCCGAGGGAACGAGATAAAGATCACACA 603

184 rPcYsLeuProIstYrProLeuProGlnGlyThrGlnAspGlnIstYr 200

604 GCACGATACCCAGTTTGTCTGCCATGCTGGCGCCCTCTCTCTTGAT 653

201 AlaThrIleProSerLeuSerAlaMetLeuGlyAlaLeuPheLeuTrpMe 217

654 GTTCTGGCAGATTTCACCTGCTGTGAGAGTCAATCGAAGGA 703

217 tPheTrpProSerPheAsnSerAlaLeuLeuArgSerProIleGlnArgL 234

704 AGAATGCCGTGTACACACTACTATGCTGTAGCAGTCAAGCTGTGACA 753

234 ysaAsnAlaValAlaPheAsnThrTyrlAlaValAlaValSerValIstYr 250

754 GCCATCTCAGGGTATCTTGGCTCACCCCAAGGGAAGATCACAGAC 803


```

251 AlaIleSerGlySerSerLeuAlaHisProGlnGlyLysIleSerLysTh 267
804 TTATGTGCACAGTGGCTGTGGTGGCAGAGCGGTGGCTGGTACCTGCT 853
|||||
267 rTyValHisSerAlaValIleAlaGlyGlyValAlaValAlaGlyThrSec 284
884 GTACACCTATCCCTTCTCCGCTGGCTGGCCATGGGCTGGCTGGCTGCT 903
|||||
284 ySHISleuIleProSerProTrpLeuAlaMetValIleuGlyLeuValAla 300
904 GGGCGATCTCCGCTGGGGGAGCCAGTACCTCCGGGGGTGGTGAACCG 953
|||||
301 GlyLeuIleSerValGlyGlyAlaLysTyrLeuProGlyLysCysAsnAr 317
954 AGTGTGGGATTTCCCGACAGCTCCATCATGGGCTACAACTTCACCTTGC 1003
|||||
317 gValIleuGlyIleProHisSerSerIleMetGlyTyrAsnPheserIle 334
1004 TGGGTCTGCTGGAGAGATCATCTACATTTGCTGCTGCTGCTTATAC 1053
334 euGlyLeuIleuGlyIleIleTyrIleValIleuLeuValIleuAspTr 350
1054 GTGGGAGCGCGCATGTCATGTCCTCCAGGCTCCCTCAGCATTCG 1103
|||||
351 ValGlyAlaGlyAsnGlyMetIleGlyPheGlnValIleuLeuSerIleG 367
1104 GGAATCAGCTGGCCATCGTATAGCTCTCAGCTGCTGGCTCCAGACG 1153
|||||
367 yGluLeuSerIleuAlaIleValIleAlaIleuThrSerGlyLeuLeuTh 384
1154 GTTGTCTCTTAATCTTAATAATATGAAAGCCTCATGAGGCTTAATAT 1203
384 yLeuLeuLeuAsnLeuLysIleTrpLysAlaProHisGlnAlaLysTyr 400
1204 TTTGTATGACCAAGTTTCTGGAAGTTCTCATTTGGCTGGATTT 1251
|||||
401 PheAspAspGlnValPheTrpLysPheProHisLeuAlaValGlyPhe 416
seq_name: SwissProt_39:RHLR_PANTR

seq_documentation_block:
ID RHLR_PANTR STANDARD; PRT; 416 AA.
AC Q28B14;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DE RH-LIKE PROTEIN IIR (RHESUS-LIKE PROTEIN IIR).
OS Pan troglodytes (Chimpanzee).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Pan.
OX NCBI_TaxID=9598;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Bone marrow;
RX MEDLINE=95083595; PubMed=7993375;
RA Salvignol I., Blancher A., Calvas P., Clayton J., Socha W.W.,
RA Colin Y., Ruffie J.;
RT "Molecular genetics of chimpanzee Rh-related genes: their
RT relationship with the R-C-E-F blood group system, the chimpanzee
RT counterpart of the human rhesus system.";
RT Biochem. Genet. 32:201-221(1994).
RL -|- FUNCTION: MAY BE PART OF AN OLIGOMERIC COMPLEX WHICH IS LIKELY TO
CC HAVE A TRANSPORT OR CHANNEL FUNCTION IN THE ERYTHROCYTE MEMBRANE.
CC -|- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -|- SIMILARITY: BELONGS TO THE RH FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

```

```

CC -----
DR EMBL: L37050; AAA65624.1; -.
DR InterPro: IPR001905; -.
DR InterPro: IPR002229; -.
DR Pfam: PF00909; Ammonium transp; 1.
DR PRINTS: PR00342; RHESUSRHD.
KW Erythrocyte; Transmembrane.
FT INIT_MET 0
FT TRANSMEM 11 31
FT TRANSMEM 43 63
FT TRANSMEM 76 96
FT TRANSMEM 124 144
FT TRANSMEM 171 191
FT TRANSMEM 202 222
FT TRANSMEM 237 257
FT TRANSMEM 264 284
FT TRANSMEM 286 306
FT TRANSMEM 330 350
FT TRANSMEM 357 377
SQ SEQUENCE 416 AA; 45299 MW; C055D7CC8BA0420 CRC64;

```

```

alignment_scores:
Quality: 1970.00
Ratio: 4.888
Percent Similarity: 96.875
Percent Identity: 93.029

```

alignment_block:

US-09-600-714-41 x RHLR_PANTR ..

Align seg 1/1 to: RHLR_PANTR from: 1 to: 416

```

4 ACCTTAAGTACCCGGGCTGTCCGGGCTGCTGCCCCCTGGGCGCT 53
|||||
1 SerSerLysTyrProArgSerValAlaArgGlySerProLeuGlyAlaIle 17
54 AACACTGGAAGCAGCTTCATCTCCCTCTCTATTTTACCACACTATG 103
|||||
17 uThrLeuGlnAlaAlaLeuIleLeuLeuPheTyrPhePheThrGlnTyrA 34
104 AGCTTTCTTAGAGATCAAAAGGGCTGCTGGCATCCATCAAGTTGGC 153
|||||
34 sPAlaSerLeuGlnAspGlnLysGlyLeuValAlaSerTyrGlnValGly 50
154 CAAGATTCAGACCGTGTATGGCGGCAFTTGCTGGGTTCCACCTCGAG 203
|||||
51 GlnAspLeuThrValMetAlaAlaIleGlyPheGlyPheLeuThrSerSe 67
204 TTTCGGAGACACAGCTGAGCAGTGTGGCTTCAACCTCTTCATGCTGG 253
|||||
67 rPheArgAlaHisSerTrpSerSerValAlaPheSerLeuPheMetLeuA 84
254 CGCTTGCTGTGAGTGGGCAATCCCTGCGAGGGCTTCTGAGCCAGTTC 303
|||||
84 lAlaLeuGlyAlaGlnTrpAlaIleLeuLeuAspGlyPheLeuSerGlnPhe 100
304 CCTTGGGGAAGGTGATCATCACATGTTCACTATTCGGCTGGCCACCAT 353
|||||
101 ProProGlyLysValValIleThrIleuPheSerIleArgGlyAlaThrTh 117
354 GAGTCCCTTGTGGTGTGATCTCAAGTGAGTCTGTGGGGAAGGTCA 403
|||||
117 rSerAlaLeuSerValIleuIleSerValAspAlaValIleuGlyValA 134
404 ACTTGCGCAGTGTGGTGTGATGCTGCGTGGAGTGCACCTTAGGC 453
|||||
134 sIleuValGlnIleuValValMetValIleuValGlnValThrAlaLeuGly 150
454 AACCTGAGATGTCATCATGATATCTTCAACAGACAGTACCATGAA 503
|||||
151 ThrValArgMetValIleSerAsnIlePheAsnTrpAspTyrHisMetAs 167
504 CATGATGCACATCTACGTTGTGCAAGCCATATTTGGGCTGTGTGGCCT 553

```

```

167 nleuethistiletyValpnealaatyrPheyleuSerValaIat 184
554 GGTGGCTGCCAAAGCCTTACCCGGAGGAGGATGAAGATGACGCA 603
184 rPcysleuProlyserProleupPolyglythrGluaspLysaspGlnle 200
604 GCAACGATACCAGTTGTCTGCGCATGCTGGCGGCCCTCTCTTGAT 653
201 AlaThrIleProserleuSerAlaMetleuGlyAlaLeuPheleuTrpme 217
654 GTTCGGCCAAAGTTCAACTGCTGCTCTGAGAAAGTCCAAATCGAAGA 703
217 tPheThrProSerPheasnSerAlaLeuLeuArgSerProIleGluArgL 234
704 AGAAGCCGCTGTTCACACCTACTAGCTGTAGCGTAGCGGTGAGCA 753
234 ySasAlaValAlaPheasnThrTyrrAlaValAlaValaSerValaThr 250
754 GCCATCTGAGGATCATCCTTGCTCACCCCGACAGGAGATCAGCAAGAC 803
251 AlaIleSerGlySerSerleuAlaHisProGlnGlyLysIleSerMetse 267
804 TTAATGTCACAGTGGCGGTGTGGCAGAGCGCTGGCTGGCTGCTGCT 853
267 rTyMetHisasnaAlaValleuAlaGlyAlaValaGlyThrSerc 284
854 GTACACCTATCCCTCTGCTGGCTGCTGCTGCTGCTGCTGCTGCTGCT 903
284 ySHisLeuIleProserProItrPleuAlaMetValleuGlyLeuVala 300
904 GGGCTGATCTCCGTCGGGGGAGCCAACTAGTCCGGGGGTGTGTACCG 953
301 GlyleuIleSerValGlyAlaValaLysTyrrleuProGlyCysAsnAr 317
954 AGTGTGGGGATGCCACAGCTCATCATGGCTGACAACTTCAGCTTGC 1003
317 gValleuGlyIleProHisSerSerValMetGlySerAsnPheSerTrpL 334
1004 TGGGTCTGCTGGAGAGATCATCTACATTTGCTGCTGCTGCTGCTGAT 1053
334 euGlyLeuLeuGlyGluIleIleTyrrIleValleuValaArgHisThr 350
1054 GTCCGAGCCGCAATGSCATGATGGCTTCAGGCTCTCTCAGCATTTGG 1103
351 IleTrpaenGlyAsnGlyMetCileGlyPheGlnValleuAlaArgIleG 367
1104 GGAACCTGAGCTTGGCCATCGATAGCTTCACGCTGCTGCTCCGACAG 1153
367 yGluPheSerleuAlaThrTrhIleAlaLeuThrSerGlyLeuLeuThrG 384
1154 GTTGGCTCTTAATCTTAAATATGGAAGACCTCATGAGGCTTAATAT 1203
384 lyleuLeuLeuAsnleuLysIleTrpLysAlaPronHisGlnAlaLysTy 400
1204 TTTGATGACCAAGTTTCTGGAAGTTTCTCATTTGGCTGTGGATT 1251
401 PheaspaspArgInValPheTrpLysPheThrHisLeuAlaValGlyPhe 416
seq_name: SwissProt_39:RHLD_GORGO
seq_documentation_block:
ID RHLD_GORGO STANDARD; PRT; 416 AA.
AC 028427;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DE RHESUS-LIKE PROTEIN.
OS Gorilla gorilla gorilla (Lowland gorilla).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Gorilla.
OX NCBI_TaxID=9595;
RN [1]

```

```

RP SEQUENCE FROM N.A.
RC TISSUE-Bone marrow;
RX MEDLINE=95085595; Pubmed=7993375;
RA Salvagnol I., Blancher A., Calvas P., Clayton J., Socha W.W.,
RA Colin Y., Ruffie J.;
RT "Molecular genetics of chimpanzee Rh-related genes: their
RT relationship with the R-C-E-F blood group system, the chimpanzee
RT counterpart of the human rhesus system.";
RL Biochem. Genet. 32:201-221(1994).
CC -!- FUNCTION: MAY BE PART OF AN OLIGOMERIC COMPLEX WHICH IS LIKELY TO
CC HAVE A TRANSPORT OR CHANNEL FUNCTION IN THE ERYTHROCYTE MEMBRANE.
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -!- SIMILARITY: BELONGS TO THE RH FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sdb.ch/announce/
CC or send an email to license@sdb.ch).
CC -----
DR EMBL: L37053; AA65627.1;
DR InterPro: IPR001905;
DR Pfam: PF00909; Ammonium_transp. 1.
DR PRINTS: PR00342; RHESUSRD.
DR KMW Erythrocyte; Transmembrane.
FT INIT MET 0
FT TRANSMEM 11 31 BY SIMILARITY.
FT TRANSMEM 43 63 POTENTIAL.
FT TRANSMEM 76 96 POTENTIAL.
FT TRANSMEM 124 144 POTENTIAL.
FT TRANSMEM 171 191 POTENTIAL.
FT TRANSMEM 202 222 POTENTIAL.
FT TRANSMEM 237 257 POTENTIAL.
FT TRANSMEM 264 284 POTENTIAL.
FT TRANSMEM 286 306 POTENTIAL.
FT TRANSMEM 330 350 POTENTIAL.
FT TRANSMEM 357 377 POTENTIAL.
SQ SEQUENCE 416 AA; 44838 MW; 398877E2B7BEE802 CRC64;

```

```

alignment_scores:
  Quality: 1948.00      Length: 416
  Ratio: 4.870         Gaps: 0
  Percent Similarity: 96.154  Percent Identity: 92.788

```

alignment_block:

US-09-600-714-41 x RHLD_GORGO ..

Align seg 1/7 to: RHLD_GORGO from: 1 to: 416

```

4 AGCTCTAAGTACCCGGGCTGTGTCGGCGCTGCTGCCCTCTGGGCT 53
|||||
1 SerSerLysTyrrProArgSerValaArgCysLysLeuProLeuCysAlaLe 17
54 AACACTGGAACGACCTGATCTCTCTCTTATTTTTCACCATG 103
|||||
17 uThrLeuGlnAlaAlaLeuThrLeuPheTyrrPhePheThrHisTyra 34
104 ACAGCTTCCTTAGAGATCAAAAGGGGCTGTCGATCCCTATCAAGTTGGC 153
|||||
34 spAlaSerleuGluaspGlnLysGlyLeuValaAlaSerTyrrGlnValGly 50
154 CAAGATCTGACCGTGATGGCGGCATTGGCTTGGGCTTCTCAGCTGAG 203
|||||
51 GlnaspLeuThrValaMetAlaAlaIleGlyPheGlyPheLeuThrSerse 67
204 TTTCCGGAGACACAGCTGAGACAGTGGGCTTCAACCTCTCATGTGCG 253
|||||
67 rPheArgGlyHisSerTrpSerSerValaAlaPheasnLeuPheMetLeuA 84

```



```

4 ACCTTAAGTACCGCGGTCTGTGCGGCGCTGCGCCCTCTGGGCGCT 53
   |||
1 SerSerIstYrProArgSerValArgArgCysLeuProLeuGlyAlaLe 17
54 AACACTGAGACGCTCATCTCTCTCTCTCTCTCTCTCTCTCTCTCT 103
   |||
17 uThrLeuGlnAlaAlaLeuLeuLeuPheThrPheThrGlnTyrA 34
104 ACGCTTCCCTAGAGATCAAAAGGGCTGTCGATCCCTATCAAGTTG 153
   |||
34 sPalSerLeuGlnuSpGlnuSgLyLeuValAlaSerTyrGlnValGly 50
154 CAAGATCTGACCGTGAAGGGCGCATTTGGCTTCCCTCCACCTGAG 203
   |||
51 GlnAspLeuThrValMetAlaAlaLeuGlyPheGlyPheLeuThrSe 67
204 TTTCCGGAGACACAGCTGAGACAGTGTGCGCTTCAACCTCTTCAT 253
   |||
67 rPheArgThrHisSerTyrSerSerValAlaPheSerLeuPheMetLeuA 84
254 CGCTTGTGTGAGTGGCAATCTGCTGAGCGCTTCTGAGCCAGTTTC 303
   |||
84 lAlaGlyValGlnTrpAlaIleLeuLeuAspGlyPheLeuSerInPhe 100
304 CCTTGTGGGAAGTGGTCAACACTGTTCAAGTATTTGGCTGGCCAG 353
   |||
101 ProProGlyValValIleThrLeuPheSerIleArgLeuAlaThrTh 117
354 GAGTCTTTCGGTGGTGCATCTGATCTGATGCTCTCTTGGGGAAGTC 403
   |||
117 rSerAlaLeuSerValIleLeuSerValAspAlaValLeuGlyValA 134
404 ACTTGGCCAGTGTGTGATGATGCTGCTGAGAGTGCACAGCTTTC 453
   |||
134 snLeuValGlnLeuValValMetValLeuValGlnValThrAlaLeu 150
454 AACCTGAGGATGTCATCATGTAATATCTTCAACACAGTACCACAT 503
   |||
151 ThrValArgMetValIleSerAsnIlePheAsnThrAspTyrHisMe 167
504 CATGATGACATCTACGTTTCGCGAGCTATTTGGCTGTCTGGGCG 553
   |||
167 nLeuMetHisIleTyrValPheAlaAlaTyrPheGlyLeuSerVal 184
554 GTTGGCTGCCAAAGCCTTACCCGAGGAGACGAGATTAAGATCAGA 603
   |||
184 rPcysLeuProIysProLeuProIysGlyThrGlnuSpLysAspGln 200
604 GCAACGATACCCAGTTTGTCCGATGCTGGCGCGCTTCTCTTGTG 653
201 AlaThrIleProSerLeuSerAlaMetLeuGlyAlaLeuPheLeu 217
654 GTTCTGGCGCAAGTTTCACTGCTGCTCTGAGAGTCCATTCAGAA 703
217 lPheThrProSerPheAsnSerAlaLeuLeuArgSerProIleGlnu 234
704 AGAATGCGGTTCACACACTACTAGTCTGAGAGTCCAGCGTGTGAC 753
234 ySAsnAlaValPheAsnThrTyrTyrAlaValAlaValSerValThr 250
754 GCCATCTCAGGGTATCTTGGCTACCCCGAAGGAGATCAGCAAGAC 803
251 AlaIleSerGlySerSerLeuAlaHisProGlnuSgLyLysIleSe 267
804 TTAATGTGACAGTCCGCTGTGCGAGAGGCGTGTGGTACTCTGCT 853
267 rTyrMetHisAsnAlaValLeuAlaGlyGlyValAlaValAlaGly 284
854 GTCACTGATCCCTTTCCTGCTGCTGCGAGGCGTGGCTGTGGCT 903
284 ySHisLeuIleThrSerProThrPheAlaMetValLeuGlyLeuVal 300

```

```

904 GGGCTGATCTCCGTCGGGGGAGCCAGTACCTGCCGGGCTGTGAACG 953
   |||
301 GlnLeuIleSerIleGlyAlaAluStryLeuProGlyCysCysAsnA 317
954 AGTGGTGGGATTCGCCACAGCTTCATCATGGGCTACAACTTCAG 1003
   |||
317 vAlaLeuGlyIleTyrHisSerSerValMetHisTyrAsnPheSer 334
1004 TGGGCTGCTTGGAGACATCATCTACATTTGCTGCTGCTGATPAC 1053
   |||
334 euGlyLeuLeuGlyGlnIleIleTyrIleValLeuLeuValHisHis 350
1054 GTCGGAGCCGCAATGCGATGATTTGGCTTCCAGTCTCTCACCAT 1103
   |||
351 ValTrpAsnGlyAsnGlyMetIleGlyPheGlnValLeuLeuArg 367
1104 GGAATCTGAGCTTGGCCATCGATAGCTCTACGCTGTGCTCTGAC 1153
   |||
367 yGlnPheSerLeuAlaThrThrIleAlaLeuThrSerIleLeuThr 384
1154 GTTGTCTCTTAATCTTAAATATGAAAGACCTCATGAGCTAAAT 1203
   |||
384 lYleuLeuLeuAsnLeuLysIleTyrPylsalapronHisAlaAla 400
1204 TTTGATGACCAAGTTTCTGGAAGTTTCTCATTTGGCTGTGATTT 1251
401 PheAspAspGlnValPheTrpLysPheProHisLeuAlaGluPhe 416

```

seq_name: SwissProt_39: RHCE_HUMAN

seq_documentation_block:

ID RHCE_HUMAN STANDARD; PRT; 416 AA.

AC P18577; Q02163; Q02164; Q02165; Q16160;

DT 01-NOV-1990 (Rel. 16, Created)

DT 01-NOV-1990 (Rel. 16, Last sequence update)

DT 01-OCT-2000 (Rel. 40, Last annotation update)

DE BLOOD GROUP RH(Ce) POLYPEPTIDE (RHESUS C/E ANTIGENS) (RH30A) (RH1XB)

DE (RH POLYPEPTIDE 1) (RHP1).

GN RHCE OR RHC OR RHE.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.

RC TISSUE=Bone marrow;

RX MEDLINE=90349591; PubMed=1696722;

RA Cherif-Zahar B., Bloy C., le van Kim C., Blanchard D., Bailly P.,

RA Hermand P., Salmon C., Carton J.-P., Collin Y.;

RT "Molecular cloning and protein structure of a human blood group Rh

RT polypeptide.";

RL Proc. Natl. Acad. Sci. U.S.A. 87:6243-6247(1990).

RN [2]

RP SEQUENCE FROM N.A.

RX MEDLINE=91058522; PubMed=2123099;

RA Avenet N.D., Ridgwell K., Tanner M.J.A., Anstee D.J.;

RT "CDNA cloning of a 30 kDa erythrocyte membrane protein associated

RT with Rh (Rhesus)-blood-group-antigen expression.";

RL Biochem. J. 271:821-825(1990).

RN [3]

RP SEQUENCE FROM N.A.

RX MEDLINE=93216282; PubMed=7916743;

RA Kajii E., Umenishi F., Iwamoto S., Ikemoto S.;

RT "Isolation of a new cDNA clone encoding an Rh polypeptide associated

RT with the Rh blood group system.";

RL Hum. Genet. 91:157-162(1993).

RN [4]

RP SEQUENCE FROM N.A. (FORMS RHIV, RHVI AND RHVII).

RC TISSUE=Bone marrow;

RX MEDLINE=92360855; PubMed=1379850;

RA le van Kim C., Cherif-Zahar B., Raynal V., Mouto I., Lopez M.,

RA Carton J.-P., Collin Y.;

RT "Multiple Rh messenger RNA isoforms are produced by alternative


```

204 TTTCGGAGACACAGCTGGAGACAGTGTGGCCTTCAACCTTTCATGCTGG 253
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
67 nPheArGhrlsSerTrpSerSerValAlaPheAsnLeuPheMetLeuA 84
254 CGCTTGGTGTGACAGTGGCAATCCTGCTGGACGGCTTCTTCAAGCAGTTC 303
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
84 lAeuGlyValGlnTrpAlaIleLeuLeuAspGlyPheLeuSerGlnPhe 100
304 CCTTGTGGAGAGTGGTGCATCAGCTGTTCAGTATTCGGCTGGCCACCAT 353
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
101 ProProGlyLysValValIleThrLeuPheSerIleArgGlyLeuAlaThrMe 117
354 GAGTGCCTTTGCGGTGATCTCAGTGAGATCTGTCTTGGGGAGAGTCA 403
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
117 tSerAlaMetSerValIleLeuIleSerAlaGlyAlaValLeuGlyLysValA 134
404 ACTTGGCCCACTGGTGTGATGGTGTGGTGGAGGTGACAGCTTTAGGC 453
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
134 snLeuAlaGlnLeuValValMetValLeuValGlyValThrAlaLeuGly 150
454 AACCTGAGAGTGCATCAGTAAATATCTTCAACAGAGCTACCCACATGAA 503
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
151 ThrLeuArgMetValIleSerAsnIlePheAsnThrAspIYrHisMetAs 167
504 CATGATGCACATCTACGTGTTCGAGAGCCTATTGGGCTGTGTGGCCT 553
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
167 nLeuArgHisPheYrValPheAlaAlaTrpPheGlyLeuThrValAlaTr 184
554 GGTGCTGCCCAAGCCTCTACCCGAGGAGACGAGAGATTAAGATCAGACA 603
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
184 rPcysLeuProLysProLeuProLysGlyThrGlnAspAspAspGlnArg 200
604 GCAACGATACCAGCTTGTGTCCATGCTGGCGGCCCTCTCTGTGGAT 653
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
201 AlaThrIleProSerLeuSerAlaMetLeuGlyAlaLeuPheLeuTrpMe 217
654 GTTCTGGCAAGTTTCACTGCTGCTGCTGAGAAGTCCAAATCGAAGA 703
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
217 tPheTrpProSerValAsnSerProLeuLeuArgSerProIleGlnArgL 234
704 AGAATGCCGCTTTCACACACTACTATGCTGTAGACAGTACGGGTGTGCA 753
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
234 ySaAsnAlaMetPheAsnThrYrYrAlaLeuAlaValSerValValThr 250
754 GGCATCTCAGAGGTCATCCTTGGCTACCCCAAGGAGATCAGACAAC 803
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
251 AlaIleSerGlySerSerLeuAlaHisProGlnArgLysIleSerMetH 267
804 TTATGTGCACAGTGCAGTGTGGAGAGGCGTGTGGTACCTGCT 853
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
267 rTyValHisSerAlaValLeuAlaGlyGlyValAlaValAlaIlyThrSerC 284
854 GTCAACCTGATCCCTTCCGCTGGCTGGCATGCTGCTGGCTGTGGCT 903
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
284 yShsIleuIleProSerProTrpLeuAlaMetValLeuGlyLeuValAla 300
904 GGGCTGATCTCCGCTGGGGAGCAAGTACTCGCGGGGTGTGTAAACCG 953
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
301 GlyLeuIleSerIleGlyGlyAlaIlyCysLeuProValIcysCysAsnArg 317
954 AGTGCCTGGGATCCCAACAGCTCCATCATGGGCTTACAACTTCAAGCTTGC 1003
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
317 gValLeuGlyIleHisIleSerValMetHisSerIlePheSerLeuL 334
1004 TGGGTCTGCTGGAGAGATCATCTACATTCGTGCTGCTGCTGATACC 1053
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
334 euGlyLeuLeuGlyGlyIleThrYrIleValLeuLeuValLeuHisThr 350
1054 GTTCGAGCGCGCAATGCGATGATGGCTTCAGATCCCTCAGACATGG 1103
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
351 ValTrpAsnGlyAsnGlyMetIleGlyPheGlnValLeuLeuSerIleG 367

```

```

1104 GGAACGACGCTTGGCCATCGTATAGCTCTCACGCTGTCTCTCGACAG 1153
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
367 yGluLeuSerLeuAlaIleValIleAlaLeuThrSerGlyLeuLeuThrG 384
1154 GTTGTCTCTTAATCTTAAATATGGAAGCAGCTCATGAGCTTAATAT 1203
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
384 lLeuLeuLeuAsnLeuLysIleTrpLysAlaProHisValAlaLysTyr 400
1204 TTTGATGACCAAGTTTCTGGAGCTTCCATTCATTTGGCTGTGGATTT 1251
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
401 PheAspAspGlnValPheTrpLysPheProHisLeuAlaValGlyPhe 416

```

Seq_name: SwissProt_39:RHLE_PANTR

seq_documentation_block:

ID RHLE_PANTR STANDARD; PRT; 416 AA.

AC 028812;

DT 01-NOV-1997 (Rel. 35, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)

DT 01-NOV-1997 (Rel. 35, Last annotation update)

DE RH- LIKE PROTEIN IIF (RHESUS- LIKE PROTEIN IIF).

OS Pan troglodytes (Chimpanzee).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Pan.

OX NCBI_TaxID=9598;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=9505595; PubMed=7993375;

RA Salvignol I., Blancher A., Calvas P., Clayton J., Socha W.W.,

RA Colin Y., Ruffie J.;

RT "Molecular genetics of chimpanzee Rh-related genes: their

RT relationship with the R-C-E-F blood group system, the chimpanzee

RT counterpart of the human rhesus system.";

RL Biochem. Genet. 32:201-221(1994).

CC - FUNCTION: MAY BE PART OF AN OLIGOMERIC COMPLEX WHICH IS LIKELY TO

CC HAVE A TRANSPORT OR CHANNEL FUNCTION IN THE ERYTHROCYTE MEMBRANE.

CC - SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.

CC - SIMILARITY: BELONGS TO THE RH FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)CC or send an email to license@sib-sib.ch).

CC -----

DR EMBL: L37048; AAA5622.1; -

DR InterPro: IPR001905; -

DR InterPro: IPR002229; -

DR Pfam: PF00909; Ammonium_transp. 1.

DR PRINTS: PR00342; RHESUSRHD.

KW Erythrocyte; Transmembrane.

FT INIT_MET 0 BY SIMILARITY.

FT TRANSMEM 11 31 POTENTIAL.

FT TRANSMEM 43 63 POTENTIAL.

FT TRANSMEM 76 96 POTENTIAL.

FT TRANSMEM 124 144 POTENTIAL.

FT TRANSMEM 171 191 POTENTIAL.

FT TRANSMEM 202 222 POTENTIAL.

FT TRANSMEM 237 257 POTENTIAL.

FT TRANSMEM 264 284 POTENTIAL.

FT TRANSMEM 286 306 POTENTIAL.

FT TRANSMEM 330 350 POTENTIAL.

FT TRANSMEM 357 377 POTENTIAL.

SQ SEQUENCE 416 AA; 45136 MW; 325ED916268BAF2F CRC64;

alignment_scores:

Quality: 1914.00 Length: 416

Ratio: 4.833 Gaps: 0

Percent Similarity: 95.192 Percent Identity: 91.106

alignment_block:
US-09-600-714-41 x RHLE_PANTR ..

Align seg 1/1 to: RHLE_PANTR from: 1 to: 416

```

4 AGCTTAAGTACCGCGGTGTCTCCGGCGCTCCGCCCTCTGGCCCT 53
1 SerSerLysTyrProArgSerValArgCysLeuProLeuGlyAla 17
54 AACACTGGAAGAGCTCATCTCTCTCTCTCTCTCTCTCTCTCTCT 103
17 uThrLeuGluAlaAlaLeuLeuLeuLeuPheTyrPhePheThrHisTyrA 34
104 ACGCTTCTTAGAGGATCAAAAGGGGCTCTGGCATCTCATCAAGTTGC 153
34 spAlaSerLeuGlnuSpGlnuGlyLeuValAlaSerTyrGlnValGly 50
154 CAAGATCTGACCGTGATGGCGCATTTGGCTTGGCTTCTCAGCTGAG 203
51 GlnAspLeuThrValMetAlaAlaIleGlyPheGlyPheLeuThrSerSe 67
204 TTTCGGAGACACAGCTGAGCAGTGGGCTTCAACCTTCATGCTGG 253
67 rPheArgAlaHisSerTyrPheSerValAlaPheAsnLeuPheMetLeuA 84
254 CGCTTGTGTGCAGTGGCGCATCTCTGTGAGCGCTTCTCAGCCAGTTG 303
84 lAlaGlyValGlnTrpAlaIleLeuLeuAspGlyPheLeuSerGlnPhe 100
304 CCTTCTGGGAAGGTGTATCATCACATCTTCATTCCTGGCTGCCACCAT 353
101 ProProGlyLysValValIleThrLeuPheSerIleArgLeuAlaThrH 117
354 GAGTCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 403
117 rSerAlaLeuSerValLeuIleSerAlaGlyAlaValLeuGlyTyrValA 134
404 ACTTGCGCGCAGTTGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 453
134 snLeuValGlnLeuValValMetValLeuValGlnValThrAlaLeuGly 150
454 AACCTGAGATGTGTATCATCTTATCTTCAACACAGATACCATGGA 503
151 ThrMetArgMetValIleSerAsnIlePheAsnThrAspTyrHisMetAs 167
504 CAGTATGCATCTAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 553
167 mLeuThrHisPheTyrLeuPheThrAlaTyrPheGlyValIleThrValAla 184
554 GGTGCTGCGCAAGCCTTACCGCAGGAGGAGGAGGATAAAGATCAGACA 603
184 rPcysLeuProLysProLeuProAspValIleGlnuSpLysAspGlnIle 200
604 GCAAGATACCAAGTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 653
201 AlaThrIleProSerLeuSerAlaMetLeuGlyAlaLeuPheLeuTyrPhe 217
654 GTTTCGGCAAGTTTCACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 703
217 tPheTrpProSerPheAsnSerAlaLeuLeuArgSerProIleGlnuArgL 234
704 AGAATCCCTGTTCACACCTACTATGTGTAGCATCAGCTGAGGTGAGA 753
234 yAsnAlaValAlaPheAsnThrTyrTyrAlaLeuAlaValSerValValThr 250
754 GCCATCTAGGGTATCTTGGCTACCCCAAGGAGGAGATCAGCAAGAC 803
251 AlaIleSerGlySerSerLeuAlaHisProGlnuGlyLysIleSerMetH 267
804 TTATGTGCAGTGGCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 853
267 rTyrValHisSerAlaValLeuAlaGlyValAlaValAlaValGlyThrSerC 284

```

```

854 GTACCTGATCCCTTCTCCGGCTTGCACATGCTGCTGGCTTGTGCT 903
284 ysnIleuIleProSerProTrpLeuAlaMetValLeuGlyLeuValAla 300
904 GGGTGTATCTCCGTGGGGAGCCCAAGTACCTGCCGGGGTGTGTACCG 953
301 GlyLeuIleSerIleGlyAlaLysCysGlyProGlyCysCysAsnAr 317
954 AGTGTGGGATTTCCCAAGCTCATGAGGATGAGTACCACTCAGCTTGC 1003
317 gValLeuGlyIleProAspSerSerValMetHisTyrAsnPheSerLeu 334
1004 TGGTGTCTGTGAGAGATCATCTACATTTGCTCTGCTGCTGATACC 1053
334 euGlyLeuLeuGlyIleuIleTyrIleValLeuValAlaArgHisThr 350
1054 GTTCGAGCCGGCAATGAGCATGATTTGGCTTCCAGGCTTCCTAGCATGG 1103
351 ValTrpAsnGlyAsnGlyMetIleGlyPheGlnValLeuLeuSerMetC 367
1104 GGAACCTCAGCTTGGCAGTCATGATAGCTCTCAGCTGTGCTCTGACAG 1153
367 yGluLeuSerLeuAlaIleAlaIleAlaLeuThrSerGlyLeuLeuThg 384
1154 GTTTCCTCTTAATCTTAAATATGGAAGACCTCATGAGGCTAAATAT 1203
384 lYleuLeuLeuAsnLeuLysIleTrpLysAlaProHisValAlaLysTyr 400
1204 TTGTGAGCAAGTTTCTGGAAGTTTCTCATTTGGCTTGGCTGAGATT 1251
401 SerAspAspGlnValAlaPheTrpLysPheProHisLeuAlaValGlyPhe 416

seq_name: SwissProt_39:RHLC_GORGO
seq_documentation_block:
ID RHLC_GORGO STANDARD; PRT; 416 AA.
AC Q28426;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE RH-LIKE PROTEIN IC (RHESUS-LIKE PROMEIN IC).
OS Gorilla gorilla gorilla (Lowland gorilla).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Gorilla.
OX NCBI_TaxID=9595;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Bone marrow;
RX MEDLINE=95085595; PubMed=7993375;
RA Salvignol I., Blancher A., Calvas P., Clayton J., Socha W.W.,
RA Collin Y., Ruffie J.;
RT "Molecular genetics of chimpanzee Rh-related genes: their
RT relationship with the R-C-E-F blood group system, the chimpanzee
RT counterpart of the human Rhesus system.";
RL Biochem. Genet. 32:201-221(1994).
CC -1- FUNCTION: MAY BE PART OF AN OLIGOMERIC COMPLEX WHICH IS LIKELY TO
CC HAVE A TRANSPORT OR CHANNEL FUNCTION IN THE ERYTHROCYTE MEMBRANE.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -1- SIMILARITY: BELONGS TO THE RH FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: L37052; AAA65626.1; -
DR InterPro: IPR001905; -
DR InterPro: IPR002229; -
DR Pfam: PF00909; AminoLium_transp; 1.
DR PRINTS: PR00342; RHESUSRHD.

```


KW Erythrocyte; Transmembrane.
 FT INTL MET 0 0 BY SIMILARITY.
 FT TRANSMEM 11 31 POTENTIAL.
 FT TRANSMEM 43 63 POTENTIAL.
 FT TRANSMEM 76 96 POTENTIAL.
 FT TRANSMEM 124 144 POTENTIAL.
 FT TRANSMEM 171 191 POTENTIAL.
 FT TRANSMEM 202 222 POTENTIAL.
 FT TRANSMEM 237 257 POTENTIAL.
 FT TRANSMEM 264 284 POTENTIAL.
 FT TRANSMEM 286 306 POTENTIAL.
 FT TRANSMEM 330 350 POTENTIAL.
 FT TRANSMEM 357 377 POTENTIAL.
 SQ SEQUENCE 416 AA; 45129 MW; 0F3B623F908E087 CRC64;

alignment_scores:
 Quality: 1903.00 Length: 416
 Ratio: 4.806 Gaps: 0
 Percent Similarity: 95.192 Percent Identity: 89.904

alignment_block:
 US-09-600-714-41 x RHLG-GORGO

Align seg 1/1 to: RHLG-GORGO from: 1 to: 416

```

4 AGCTTAAGTACCCGGCTGTGCGGCGCTGCGCCCTGCGCCCT 53
| | | | | | | | | | | | | | | | | | | | | | | |
1 SerSerLysTyrProArgSerValArgCysLeuProLeuCysAlaIe 17
| | | | | | | | | | | | | | | | | | | | | | | |
54 AACACTGGAAGCAAGCTTCATTCTCTCTTATTTTATACCACTATG 103
| | | | | | | | | | | | | | | | | | | | | | | |
17 uThrLeuGlnAlaIleuIleLeuLeuPheTyrPhePheThrHisTyrA 34
| | | | | | | | | | | | | | | | | | | | | | | |
104 AGCTCTCTTAGAGATCAAAAGGGCTGCGCATCTCATCAATGGC 153
| | | | | | | | | | | | | | | | | | | | | | | |
34 sPAlaSerLeuGlnAspGlnLysGlyLeuValAlaSerTyrGlnValGly 50
| | | | | | | | | | | | | | | | | | | | | | | |
154 CAAGATCTGACGATGAGCGGCAATGCGCTGGGCTCTCCACCTCGAG 203
| | | | | | | | | | | | | | | | | | | | | | | |
51 GlnAspLeuThrValMetAlaAlaIleGlyPheGlyPheLeuThrSerSe 67
| | | | | | | | | | | | | | | | | | | | | | | |
204 TTTCCGAGACACAGCTGAGAGAGTGTGGCTTCAACCTTCATGCTGG 253
| | | | | | | | | | | | | | | | | | | | | | | |
67 rPheArgAlaHisSerTyrSerSerValAlaPheAsnLeuPheMetLeuA 84
| | | | | | | | | | | | | | | | | | | | | | | |
254 GCGTTGGTGGAGTGGCAATCTGCTGAGAGGCTTCTGAGCCAGTTC 303
| | | | | | | | | | | | | | | | | | | | | | | |
84 lAlaGlyValGlnTyrPalaIleLeuLeuAspGlyPheLeuSerGlnPhe 100
| | | | | | | | | | | | | | | | | | | | | | | |
304 CTTCTGGGAAGTGGTGCATCACATGTTCACTATTCGGTGGCCACCAT 353
| | | | | | | | | | | | | | | | | | | | | | | |
101 ProProGlyLysValValIleThrLeuPheSerIleArgGlnAlaIleThr 117
| | | | | | | | | | | | | | | | | | | | | | | |
354 GAGTCCTTTGCGGTGATCTCAGTGAATGCTCTTGGGGAAGTCA 403
| | | | | | | | | | | | | | | | | | | | | | | |
117 tSerAlaLeuSerValIleuIleSerAlaGlyAlaValLeuGlyTyrValA 134
| | | | | | | | | | | | | | | | | | | | | | | |
404 ACTTGGCCAGTGTGGTGAATGGTGGTGGAGAGTGCATGAGCTTAAAGC 453
| | | | | | | | | | | | | | | | | | | | | | | |
134 snLeuValGlnLeuValAlaMetValIleuValGlnValThrAlaLeuGly 150
| | | | | | | | | | | | | | | | | | | | | | | |
454 AACCTGAGATGATCATGATATATCTTCAACAGACAGATCAACATGAA 503
| | | | | | | | | | | | | | | | | | | | | | | |
151 ThrMetAlaGlnMetValIleSerAsnIlePheAsnThrAspTyrHisMetAs 167
| | | | | | | | | | | | | | | | | | | | | | | |
504 CATGATGACATCTACGTTCGACGACCTATTTGGGCTGTGTGGCCT 553
| | | | | | | | | | | | | | | | | | | | | | | |
167 nMetThrHisPheTyrValPheAlaAlaTyrPheGlyValThrValAlaIat 184
| | | | | | | | | | | | | | | | | | | | | | | |
554 GGTGGCTGCCAAAGCTTACCCGAGGAGACGAGAGATAAAGATAGACA 603
| | | | | | | | | | | | | | | | | | | | | | | |
184 rPcysLeuProLysProLeuProAspIleLysGlnAspLysAspGlnIle 200

```

```

604 GCAAGATACCCAGTTTGTCGTCATGCTGGGGCGCCCTTCTTGATGAT 653
| | | | | | | | | | | | | | | | | | | | | | | |
201 AlaThrIleProSerLeuSerAlaMetLeuGlyTyrIleuPheLeuTyrPMe 217
| | | | | | | | | | | | | | | | | | | | | | | |
654 GTTTCGGCCAAAGTTTCAACTGCTGCTGCTGAGAAATCGAAAGGA 703
| | | | | | | | | | | | | | | | | | | | | | | |
217 tPheThrProSerPheAsnSerAlaLeuLeuArgSerProIleGlnAlaGly 234
| | | | | | | | | | | | | | | | | | | | | | | |
704 AGATGCGGTGTTCACACACTACTATGCTGTAGACAGTACGGTGTGACA 753
| | | | | | | | | | | | | | | | | | | | | | | |
234 yAsnAlaValAlaPheAsnThrTyrAlaLeuAlaValSerValAlaThr 250
| | | | | | | | | | | | | | | | | | | | | | | |
754 GGCATCTAGGGTTCATCTGCTGCTCACCCCAAGGAGATGACCAAGAC 803
| | | | | | | | | | | | | | | | | | | | | | | |
251 AlaIleSerValSerSerLeuAlaHisProGlnGlyLysIleAsnMetH 267
| | | | | | | | | | | | | | | | | | | | | | | |
804 TTATGTGCACAGTGGCGTGTGGCAGAGGCGTGGCTGTGGTACTCTGT 853
| | | | | | | | | | | | | | | | | | | | | | | |
267 rTyrMetHisAsnAlaValLeuAlaGlyGlyAlaValAlaGlyThrSerC 284
| | | | | | | | | | | | | | | | | | | | | | | |
854 GTCACCTGATCCCTTCGCGGCTGCTGGCATGCTGCGGCTGTGGCT 903
| | | | | | | | | | | | | | | | | | | | | | | |
284 yHisLeuIleThrSerProThrPheuAlaMetValLeuGlyLeuValAla 300
| | | | | | | | | | | | | | | | | | | | | | | |
904 GGGCTGATCTCCGTCGCGGAGACCAAGTACCTGCCGGGCTGTGTAACG 953
| | | | | | | | | | | | | | | | | | | | | | | |
301 GilyeulIleSerIleGlyGlyAlaLysCysLeuProGlyCysCysAsnAr 317
| | | | | | | | | | | | | | | | | | | | | | | |
954 AGTGTGGGGATTCGCCACAGCTCCATCATGAGGCTACACTTCAGCTTGC 1003
| | | | | | | | | | | | | | | | | | | | | | | |
317 gValIleuGlyIleHisAspSerSerValMetHisTyrAsnPheSerLeuL 334
| | | | | | | | | | | | | | | | | | | | | | | |
1004 TGGCTGCTGTGAGAGATCATCATGATTTGCTGCTGCTGCTGATACC 1053
| | | | | | | | | | | | | | | | | | | | | | | |
334 euGlyLeuLeuGlyGlyIleThrTyrIleValLeuMetValLeuHisThr 350
| | | | | | | | | | | | | | | | | | | | | | | |
1054 GTGCGAGCCGGCAATGAGCATGATTGGCTTCCAGAGCTCTCATCAATGG 1103
| | | | | | | | | | | | | | | | | | | | | | | |
351 ValGlyAlaGlnAsnGlyMetValGlyPheGlnValLeuValSerThrGl 367
| | | | | | | | | | | | | | | | | | | | | | | |
1104 GGAATCAGCTTGGCCATGATAGCTCTCAGCTGTGCTCTGACAG 1153
| | | | | | | | | | | | | | | | | | | | | | | |
367 yGluLeuSerLeuAlaLeuAlaIleAlaValThrSerGlyLeuLeuThrG 384
| | | | | | | | | | | | | | | | | | | | | | | |
1154 GTTGTCTCTTAATCTTAAATATGGAAGACCCATGAGGCTTAATAT 1203
| | | | | | | | | | | | | | | | | | | | | | | |
384 lYleuLeuLeuAsnLeuLysIleThrPlyAlaProHisAlaAlaLysTyr 400
| | | | | | | | | | | | | | | | | | | | | | | |
1204 TTTGATGACCAAGTTTCTGGAAGTTTCTCATTTGGCTGTGGATT 1251
| | | | | | | | | | | | | | | | | | | | | | | |
401 PheAspAspGlnValPheThrPlySerPheThrHisLeuAlaValGlyPhe 416

```

seq_name: SwissProt_39:RHL_MACMU

seq_documentation_block:
 ID RHL_MACMU STANDARD; PRT; 416 AA.
 AC 028849;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE RH-LIKE PROTEIN (RHEUS-LIKE PROTEIN).
 OS Macaca mulatta (Rhesus macaque).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
 OC Cercopithecoidea; Macaca.
 OC NCBTaxID=9544;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=bone marrow;
 RX MEDLINE=94223693; PubMed=7909570;
 RA Mouro I., Le van Kim C., Cherif-Zahar B., Salvignol I., Blancher A.,
 Carton J.-P., Collin Y.;

RT "Molecular characterization of the Rh-like locus and gene transcripts
from the rhesus monkey (Macaca mulatta).";
RL J. Mol. Evol. 38:169-176(1994).
CC -1- FUNCTION: MAY BE PART OF AN OLIGOMERIC COMPLEX WHICH IS LIKELY TO
HAVE A TRANSPORT OR CHANNEL FUNCTION IN THE ERYTHROCYTE MEMBRANE.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -1- SIMILARITY: BELONGS TO THE RH FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
at the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb.ch/announce/>
or send an email to license@isb.slb.ch).

CC EMBL: S70343: AAB30637.1; -.
DR InterPro: IPR001905; -.
DR InterPro: IPR002229; -.
DR Pfam: PF00909; Ammonium_transp; 1.
DR PRINTS: PR00342; RHESUSRD.
KM Erythrocyte; Transmembrane; Polymorphism.
FT INIT MET 0 0 BY SIMILARITY.
FT TRANSMEM 11 31 POTENTIAL.
FT TRANSMEM 43 63 POTENTIAL.
FT TRANSMEM 76 96 POTENTIAL.
FT TRANSMEM 124 144 POTENTIAL.
FT TRANSMEM 171 191 POTENTIAL.
FT TRANSMEM 202 222 POTENTIAL.
FT TRANSMEM 237 257 POTENTIAL.
FT TRANSMEM 264 284 POTENTIAL.
FT TRANSMEM 286 306 POTENTIAL.
FT TRANSMEM 330 350 POTENTIAL.
FT TRANSMEM 357 377 POTENTIAL.
FT VARIANT 1 55 MISSING (IN MAC-B).
FT VARIANT 70 70 A -> K (IN MAC-B).
FT VARIANT 106 106 R -> V (IN MAC-B).
FT VARIANT 191 191 P -> L (IN MAC-B).
FT VARIANT 261 261 G -> R (IN MAC-B).
FT VARIANT 342 342 Y -> C (IN MAC-B).
SQ SEQUENCE 416 AA; 45672 MW; 3BFD939BEA79AF6F CRC64;

alignment_scores:
Quality: 1642.00 Length: 416
Ratio: 4.265 Gaps: 0
Percent Similarity: 92.548 Percent Identity: 77.404

alignment_block:

US-09-600-714-41 x RHL_MACMU ..

Align seg 1/1 to: RHL_MACMU from: 1 to: 416

```

4 AGCTCTAAGTACCCGCGGTGTGTCGCGGCTGCGCCCTCTGGGCGCT 53
1 SerSerLysTYrProArgSerValAlaGlyScyLeuProLeuTrpAlaLe 17
54 AACCTGGAACGAGCTCATCTTCCCTGATATTTTAAACCCCTATG 103
17 urThreuIduaIaIaLeuIleLeuLeuPhePheThrTYrA 34
104 ACGCTTCCTAGAGATCAAAAGGGGCTGTCGATCCTATCAAGTTGGC 153
34 spAlaSerLeuGluaspGlnsglyLeuValAlaSerTYrGlnValLys 50
154 CAAGATCTGACCGTGATGGGGCCATTGGCTTGGCTTCTACCTGAG 203
51 GlnspleuThrValMetAlaValLeuGlyLeuGlyPhePheThrSerAs 67
204 TTTCGGGAGACAGACGTCGAGACGTCGCTCAACCTCTTCAAGCTGG 253
67 nleuArgIaGlnSerTYrPheSerValAlaPheAsnLeuPheLeuA 84

```

```

254 CGCTTGTGTGCAGTGGGCAATCCTGCTGAGCGCTTCTGACCACTTC 303
84 IaleuGlyValGlnTrpAlaIleLeuLeuAspGlyPheLeuSerGlnPhe 100
304 CCTCTGGGAAGGTGGTCATCAACCTGTTCAGTATTCGGCGGCACCAT 353
101 SerProGlyLysValAlaIleLysLeuPheSerIleArgLeuAlaThrAr 117
354 GAGTGTCTTGTGCGTGCATCTCAGTGGATGGATGCTGTCTTGGGGAAGTCA 403
117 gSerThrIleSerMetLeuIleSerMetAsnAlaValLeuGlyValA 134
404 ACTTGGCGCAGCTGTGTGATGTGTGCTGTGGAAGTACAGCTTTAGGC 453
134 snLeuValGlnLeuValValMetGluLeuValGlnLeuThrValPheGly 150
454 AACCTGAGCATGTCATCAATCAATATATCTTCAACACAGCTCCACATCA 503
151 ThrMetArgIleValIleAsnAsnIlePheLysIleAspTYrGlyMetAs 167
504 CATGATGCACATCTACGTTCGCGAGCCTATTTGGGCTGTGTCGCGCT 553
167 nMetMetHisIleHisValPheAlaIaIaTYrPheGlyLeuThrValAlaIa 184
554 GGTGCTGTCACAAAGCCTTACCCGAGGAGACGAGATAAAGATCAGACA 603
184 rPcysLeuProLysProLeuProLysGlyThrGluAspLysTYrGlnThr 200
604 GCACGATACCCAGTTTGTCTGCCATGCTGGCGGCCCTTCTTGTGAT 653
201 ThrThrSerProSerLeuPheAlaMetLeuGlyThrLeuPheLeuTrpMe 217
654 GTTTCGGCCAGATTTCACATGCTGCTGCTGAGAGATTCATTCGAAAGA 703
217 tPheTrpProThrPheAsnSerIleAlaLeuLeuAsnProIleGlnArgL 234
704 AGAATGCCGTGTCAACACCTACTATGCTGACGATCAGCGGTGATGACA 753
234 yAsnAlaValPheSerThrTYrAlaLeuAlaValSerAlaValThr 250
754 GCCATCTAGGGTCATCTTGGCTTCACCCCAAGGGAAGATCAGACAGAC 803
251 AlaIleSerValSerSerLeuAlaHisProGlyLysIleAsnMetTh 267
804 TTATGTGCACAGTGGGTGGTGGGAGGCGGTGCTGGGTACTGCTGCT 853
267 rTYrMetHisAsnAlaIaLeuAlaGlyGlyValAlaLeuSerAlaSerC 284
854 GTCACCTGATCCCTTCTCCGCTGGCTTGGCATGTCGTCGCTTGTGCT 903
284 ySHisValIleHisSerProTrpIleAlaMetValLeuGlyLeuValAla 300
904 GGGCTGATCTCGTGGGGGAGCCAAAGTACCTGCCGGGTGTGTAACCG 953
301 GlyLeuIleSerIleGlyGlyAlaLysCysLeuProValLysPheAsnAr 317
954 AGTGGGGGATTCGCCACAGTCATCATGAGGCTGAGCAACACTCAGCTTG 1003
317 gValLeuGlyIleHisGluSerHisSerValHisTYrThrPheGlyLeuP 334
1004 TGGGTCTGCTTGGAGAGATCATCTACATTTGCTGCTGGCTGTGATAC 1053
334 roAlaLeuLeuGlyGluIleThrTYrIleValIleMetAlaLeuArgVal 350
1054 GTGCGAGCCGCGCAATGATGATTTGGCTTCCAGGTCCTTCCTCAGATTGG 1103
351 ValTrpAlaSerSerAsnMetIleGlyPheGlnValLeuLeuSerThG1 367
1104 GGAACCTGAGCTTGGCCATGATAGCTTCACGCTGCTGCTCCGACAG 1153
367 yThrLeuSerIleuAlaMetAlaMetSerIleThrSerIleuLeuThrng 384
1154 GTTTCCTCTAAATCTTAATAATATGGAAGACCTCATGAGGCTAAATAT 1203

```

```

|||||
384  lyleuleuAaSnleuAylterpLysgLyProHlsValAlAlaLysTyr 400
1204 TTGATGACCAAGTTTCTGAGAGTTCCTCATTTGGCTGTGGATT 1251
|||||
401 PheAspAspGlnAlaPheTrpGlnPheProHlsleuAlaValGlyPhe 416
seq_name: SwissProt_39:RHL_MACFA

```

seq_documentation_block:

```

ID   RHL_MACFA          STANDARD;          PRT;          416 AA.
AC   Q28481;
DT   01-NOV-1997 (rel. 35, Created)
DT   01-NOV-1997 (rel. 35, Last sequence update)
DE   01-NOV-1997 (rel. 35, Last annotation update)
DE   RH-LIKE PROTEIN (RHESUS-LIKE PROTEIN).
OS   Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC   Cercopithecinae; Macaca.
OX   NCBI_TaxID=9541;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=bone marrow;
RX   MEDLINE=95085595; PubMed=7993375;
RA   Salvignol I., Blancher A., Calvas P., Clayton J., Socha W.W.,
RA   Collin Y., Ruffie J.;
RT   Molecular genetics of chimpanzee Rh-related genes: their
RT   relationship with the R-C-E-F blood group system, the chimpanzee
RT   counterpart of the human rhesus system.;
RL   Blochem. Genet. 32:201-221(1994).
CC   -1- FUNCTION: MAY BE PART OF AN OLIGOMERIC COMPLEX WHICH IS LIKELY TO
CC   HAVE A TRANSPORT OR CHANNEL FUNCTION IN THE ERYTHROCYTE MEMBRANE.
CC   -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC   -1- SIMILARITY: BELONGS TO THE RH FAMILY.
CC   -----
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----
DR   EMBL: L37054; AAA65628.1; -
DR   InterPro: IPR001905; -
DR   Pfam: PF00909; Ammonium_transp. 1.
DR   PRINTS: PR00342; RHESUSRHD.
KW   Erythrocyte; Transmembrane.
FT   INIT_MET          0
FT   TRANSMEM          11      31      POTENTIAL.
FT   TRANSMEM          43      63      POTENTIAL.
FT   TRANSMEM          76      96      POTENTIAL.
FT   TRANSMEM          124     144      POTENTIAL.
FT   TRANSMEM          171     191      POTENTIAL.
FT   TRANSMEM          202     222      POTENTIAL.
FT   TRANSMEM          237     257      POTENTIAL.
FT   TRANSMEM          264     284      POTENTIAL.
FT   TRANSMEM          286     306      POTENTIAL.
FT   TRANSMEM          330     350      POTENTIAL.
FT   TRANSMEM          357     377      POTENTIAL.
SQ   SEQUENCE          416 AA; 45877 MW; 17832A558D2299A9 CRC64;

```

alignment_scores:

```

Quality: 1634.00      Length: 416
Ratio: 4.311          Gaps: 0
Percent Similarity: 91.106      Percent Identity: 78.125

```

alignment_block:

US-09-600-714-41 x RHL_MACFA ..

Align seg 1/1 to: RHL_MACFA from: 1 to: 416

```

4 AGCTCTAGATACCGCGGCTGTCGCGGCGCTGCGCCCTCGGGCCCT 53
|||||
1 SerSerLysTyrTrpProHlsSerValAlaArgCysLysLeuProLeuTrpAlaLe 17
|||||
54 AACACTGGAAGACAGCTCTCATTTCTCTCTCTATTTTACCACATAG 103
|||||
17 uThrLeuGlnAlaAlaLeuIleLeuLeuPhePhePheThrTyrTyrAla 34
|||||
104 ACGCTTCTTAGAGATCAAAAGGCGCTCGTGGCATCTCATCAAGTTGGC 153
|||||
34 spAlaSerLeuGlnAlaSpGlnLysGlnLeuValAlaSerTyrGlnValCys 50
|||||
154 CAAGATCTGACCGTGTGATGGCGGCGCATGGCTTGGGCTTCCACCTCGAG 203
|||||
51 GlnAspLeuThrValMetAlaValLeuGlnLysLeuLysPhePheThrSerAs 67
|||||
204 TTTCCGGAGACACAGCTGAGACAGTGGCCCTTGACCTTCATYGGTGG 253
|||||
67 nLeuAlaArgAsnSerTrpSerSerValAlaPheAsnLeuPheLeuAla 84
|||||
254 CGCTTGGTGTGACATGGGCAATCTGCTGGAGCGGCTTCTCAGCCAGTTC 303
|||||
84 lAlaGlnValGlnTrpAlaIleLeuLeuAspGlyPheLeuSerGlnPhe 100
|||||
304 CCTTCTGGGAAGTGTGCATCATCACTGTCAGTATTCGGCTGGCCACCAT 353
|||||
101 SerProGlnLysValAlaIleLysLeuPheSerIleArgLeuAlaThrTr 117
|||||
354 GAGTGTCTTGTGCTGTGATCTCAGTGGATGCTGTGGGGAAGTCA 403
|||||
117 gSerThrThrSerMetLeuIleSerMetAsnAlaValLeuGlnLysVala 134
|||||
404 ACTTGGGCGAGTGTGGTGTGATGCTGCTGGGAGGAGACAGCTTTAGGC 453
|||||
134 snLeuAlaGlnLeuValValMetClnLeuValGlnLeuThrValPheGln 150
|||||
454 AACCTGAGGATGTCATCAGTAAATATGTTCAACACAGACTACCAATGAA 503
|||||
151 ThrMetArgIleValIleTyrAsnIlePheLysIleAspTyrGlnMetAs 167
|||||
504 CATATGCACATCTACGTTGCGCAAGCTATTTGGCGTGTGCTGGCCT 553
|||||
167 nMetMetHisIleHisValPheAlaAlaLysTyrPheGlnLeuThrValAla 184
|||||
554 GGTGCGTGGCAAGCGCTCTACCGAGGGAACGAGGATTAAGATCAGACA 603
|||||
184 TrpCysLeuProLysProLeuProLysGlnTyrGlnAspLysTyrGlnThr 200
|||||
604 GCAAGATACCCAGTTGTGTGCCATGCTGGCGCCGCTCTTCTTGGAT 653
|||||
201 ThrThrSerProSerIleuPheAlaMetLeuGlnLysThrLeuPheLeuTrpMe 217
|||||
654 GTTCTGGCAAGTTTCAACTGCTGCTGCTGAGAGAAGTCCAAATCGAAAGA 703
|||||
217 tPheTrpProThrPheAsnSerAlaLeuLeuLeuAsnProIleGlnArgL 234
|||||
704 AGAATGGCGGTGTCAACACCTACTAGCTGTAGACAGTGGGCTGTGACA 753
|||||
234 ysaAsnAlaValAlaPheSerThrTyrTyrAlaLeuAlaValSerAlaValThr 250
|||||
754 GCCATCTCAGCGTCATCTTGGCTCACCCCAAGGAAGATCAGACAGAC 803
|||||
251 AlaIleSerValIleSerSerLeuAlaHisProGlnArgLysIleAsnMetTh 267
|||||
804 TTATGTGCACAGTGGGTGTGGCAGGAGCGGTGGTGGATACCTCGT 853
|||||
267 rTyrMetProAsnAlaGlyLeuAlaGlyGlnValAlaValAlaValAlaSerC 284
|||||
854 GTCACTGATCCCTTCTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 903
|||||
284 yshIleValIleHisSerProTrpIleAlaMetValLeuGlnLysLeuValAla 300

```



```

704 AGATGCCGCTGTTCAACACCTACTATGCTGTAGCAGTACGCTGGTGCACA 753
|||||
234 ysaasnlavallphehnsrthrytyrlyalaleualavallsethrvalthr 250
|||||
754 GCCATCTCAGGGTCATCTTGGCTCACCACCAAGGAGATCAGCAAGAC 803
|||||
231 AlaileSerValserSerleualahisProginglylsileasmetlh 267
|||||
804 TTATGTGACAGTGGCGGTGGCGAGGAGGCTGGCTGGTACCTGCT 853
|||||
267 rtyrmetlissnalaValleualaglyValalavalgllythrserc 284
|||||
854 GTCACCTCCTCCCTTCCTGCTGGCTGCCATGCTGCTGGTCTGGTGGCT 903
|||||
284 yshlsleuileserSerProtrleualametvalleuvalala 300
|||||
904 GGGCTGATCTCCGTCGGGGAGACCAAGTACCTGCCGGGGTGTGTACCG 953
|||||
301 GlyleuileSerileglylValalys.....Cysleu..Pro 312
|||||
954 AGTGTGGGGATTCGCCACAGCTCATGATGGGCTACAACTTCAGCTTGC 1003
|||||
312 ..... 312
1004 TGGGTCTGCTTGGAGAGATCATCTACATTGCTGCTGCTGTGATACC 1053
|||||
312 ..... 312
1054 GTCGGAGCCGGCAATGCGATGATGGCTTCAGGCTCTCCAGCATTTGG 1103
|||||
313 .....AsptrleupProaspProleuglnhistrp 322
|||||
1104 GGAACCTAGCTTGGGACATGATGATGCTCAGCTCAGCTCTCCGACAG 1153
|||||
323 GLYTHrGlnleuglYhlsGlyAspserSerhlsValtrpPheProaspPar 339
|||||
1154 GTTGGCTCTTAATCTTAATATGGAAGACACCTCATGAGGC 1196
|||||
339 gPhehlarProlysserGlnasnmelGuserThserCysely 353
|||||
seq_name: SwissProt_39: RHAG_HUMAN
seq_documentation_block:
ID RHAG_HUMAN STANDARD: PRT: 409 AA.
AC 002094:
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE RHESUS BLOOD GROUP-ASSOCIATED GLYCOPROTEIN (ERYTHROCYTE PLASMA
DE MEMBRANE 50 KDA GLYCOPROTEIN) (RH50A).
GN RHAG.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
[1]
RP SEQUENCE FROM N.A.
RC TISSUE-Bone marrow, and Liver;
RX MEDLINE-93038558; PubMed-141776;
RA Ridgwell K., Spurr N.K., Laguda B., Macgeoch C., Avenit N.D.,
RA Tanner M.J.;
RT Isolation of cDNA clones for a 50 kDa glycoprotein of the human
RT erythrocyte membrane associated with Rh (rhesus) blood-group antigen
RT expression.
RT Biochem. J. 287:223-228(1992).
RL
RL
RP SEQUENCE OF 1-39.
RX MEDLINE-89134163; PubMed-3146980;
RA Avenit N.D., Ridgwell K., Mawby W.J., Tanner M.J.A., Anstee D.J.,
RA Kumpel B.;
RT Protein-sequence studies on Rh-related polypeptides suggest the
RT presence of at least two groups of proteins which associate in the

```

```

RT human red-cell membrane.
RL Biochem. J. 256:1043-1046(1988).
RN [3]
RP VARIANT HEMOLYTIC ANEMIA ASN-79.
RX MEDLINE-96154189; PubMed-8563755;
RA Cherif-Zahar B., Raynal V., Gane P., Mattei M.-G., Bailly P.,
RA Gibbs B., Collin Y., Cartron J.-P.;
RT "Candidate gene acting as a suppressor of the Rh locus in most cases
RT of Rh-deficiency".
RL Nat. Genet. 12:168-173(1996).
CC -1- FUNCTION: ASSOCIATED WITH RHESUS BLOOD GROUP ANTIGEN EXPRESSION.
CC MAY BE PART OF AN OLIGOMERIC COMPLEX WHICH IS LIKELY TO HAVE A
CC TRANSPORT OR CHANNEL FUNCTION IN THE ERYTHROCYTE MEMBRANE.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -1- TISSUE SPECIFICITY: ERYTHROCYTES.
CC -1- DISEASE: DEFECTS IN RHAG ARE A CAUSE OF A FORM OF CHRONIC
CC HEMOLYTIC ANEMIA ASSOCIATED WITH STOMATOCYTOSIS, AND
CC SPHEROCYTOSIS, REDUCED OSMOTIC FRAGILITY, AND INCREASED CATION
CC PERMEABILITY.
CC -1- SIMILARITY: BELONGS TO THE RH FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL: X64594; CAA45883.1;
CC DR PIR: S29124; S29124.
CC DR MIM: 180297;
CC DR MIM: 268150;
CC DR InterPro: IPR001905;
CC DR InterPro: IPR002229;
CC DR Pfam: PF00909; Ammonium_transp. 1.
CC DR PRINTS: PR00342; RHESUSRD.
CC KW Erythrocyte; Glycoprotein; Transmembrane; Disease mutation.
FT DOMAIN 1 4
FT TRANSMEM 5 24
FT DOMAIN 25 56
FT TRANSMEM 57 72
FT DOMAIN 73 83
FT TRANSMEM 84 101
FT DOMAIN 102 119
FT TRANSMEM 120 134
FT DOMAIN 135 140
FT TRANSMEM 141 153
FT DOMAIN 154 175
FT TRANSMEM 176 190
FT DOMAIN 191 208
FT TRANSMEM 209 225
FT TRANSMEM 226 238
FT TRANSMEM 239 261
FT DOMAIN 262 274
FT TRANSMEM 275 285
FT TRANSMEM 286 292
FT TRANSMEM 293 311
FT DOMAIN 312 333
FT TRANSMEM 334 351
FT DOMAIN 352 371
FT TRANSMEM 372 383
FT DOMAIN 384 409
FT CARBOHYD 37 37
FT VARIANT 79 79
FT
FT CONFLICT 2 2
FT CONFLICT 37 37
FT SEQUENCE 409 AA; 44199 MW; 1E5735FB18CB8E029 CRC64;

```

alignment_scores: 628.50

Quality: 416

Ratio: 2.131 Gaps: 9
Percent Similarity: 70.913 Percent Identity: 36.298

alignment block:

US-09-600-714-41 x RHAG_HUMAN ..

Align seg 1/1 to: RHAG_HUMAN from: 1 to: 409

```

25  GTGGGCGGCTGCTGGCCCTGGCCCTTACACCTGAGGAGGAGCTTCAT 74
    :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
1  MetArgPheThrPheProLeuMetAlaIleValLeuGluIleAlaMetCil 17
75  TCTGCTTTCTATTATTTTACCACCTAT.....GACCTTCCTTAGAG 118
    :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
17  eValLeuPheGlyLeuPheValGluTyrGluThrAspGlnThrValLeuG 34
119  ATCAAAAGGGGCTGTCGATCCATCAAGTTGGC..... 153
    :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
34  IuGlnLeuAsnIleThrLysProThrAspMetGlyIlePhePheGluLeu 50
154  .....CAAGATCTGACGCGGCGGCGGCTTGGCTTGGCTT 191
    :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
51  TyrProLeuPheGlnAspValAlaHisValMetIlePheValGlyPheGly 67
192  CCGACCTCGAGTTTCCGAGACACAGCTGAGCAGTGGCTTCAACC 241
    :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
67  eLeuMetThrPheLeuLysTyrGlyPheSerSerValGlyIleAsnL 84
242  TCTTCATGCTGGCTTGGCTTGGCAGTGGGCAATCTCTGAGCGCTTC 291
    :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
84  euleValAlaIleValLeuGlnTyrGlyThrIleValGlnGlyIle 100
292  CTGAGCCAGTCCCTTGGGAGAGTGGTCATCACATCTGACATTCG 341
    :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
101  Leu...GlnSerGlnGlyGlnLysPheAsnIleGlyIleLysAsnMetCil 116
342  GCTGGCCACCATGAGTCTTGTGGCTGCTGATCTGATCTGATGCTGTCT 391
    :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
116  eAsnAlaAspSerAlaIleAlaThrValLeuIleSerPheGlyAlaValL 133
392  TGGGGAAGTCACTGGCGCAGTGGTGGTGGTGGTGGTGGTGGTGG 441
    :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
133  euGlyLysThrSerProThrGlnMetLeuIleMetThrIleLeuGlnIle 149
442  ACAGCTTAGGCAACCTGAGTGGTCATCATCATATATCTTCAACACGA 491
    :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
150  ValPhePheAlaHisAsnGluTyrLeuValSerGluIlePheLysAlaSe 166
492  CTACACATGAAACATGATGACATCTACGTTTGGCAGCTATTTTGGGC 541
    :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
166  RasPleGlyAlaSerMetThrIleHisAlaPheGlyAlaTyrPheGlyL 183
542  TGTCTGTGGCTGGTGGCTGCCAAGCCT...CTACCCGAGGAGGAGGAG 588
    :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
183  eValAlaValAlaGlyIleLeuTyrArgSerGlyLeuArgLysLysGlu 199
589  GATAAAGATGACAGACAGACATACCCAGTTTGTGCTGCCATCTGGCGC 638
    :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
200  Asn...GluGlnSerAlaTyrTyrSerAspLeuPheAlaMetIleGlyTh 215
639  CCTCTTCTGTGGATGTTCTGGCCAGTTTCACTCTGCTGTGAGAA 688
    :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
215  rLeuPheLeuThrPheThrProSerPheAsnSerAlaIleAlaGluP 232
689  GTCCAAATGAAAGAGAAATCCGCTGTTCACACCTACTATGCTTAGCA 738
    :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
232  roGlyAspLysGlnCysArgAlaIleValAspThrTyrPheSerLeuAla 248
739  GTACGCTGGTGGTACCATCTCAGGTCATCTGCGGTGACACCCCAAG 788
    :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
249  AlacGValLeuThrAlaPheAlaPheSerSerLeuValGlnHisArgL 265
789  GAAGATCAGCAAGACTTATGTGCACAGTGGCGGTGTGGCAGAGGCGTGG 838

```

```

|||||.....  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
265  yLysLeuAsnMetValHisIleGlnAsnAlaThrLeuAlaGlyAla 282
839  CTGGGGTACCTCGTGCACCTGATCCCTTCTCGGTGGCTGACGATGGT 888
    :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
282  IValGlyThrCysAlaAspMetAlaIleHisProPheGlySerMetCile 298
889  CTGGATCTTGTGGCTGATCTCGTGCAGGAGGAGGAGGAGGAGGAGGAG 938
    :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
299  IleGlySerIleAlaGlyMetValSerValLeuGlyTyrLysPheLeuTh 315
939  GGGGTGTGTACCCAGATGCTGGGAGATTCCACACAGCTCCATCGAGCT 988
    :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
315  rProLeuPheThrThrLysLeuArgIle...HisAspThrCysGlyValH 331
989  ACAACTTC...AGCTTGTGGTCTGGTGGAGAGAGATCATCATGTTGG 1035
    :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
331  IsAsnLeuHisGlyLeuProGlyValValGlyGlyLeuAlaGlyIleVal 347
1036  CTGGTGGTCTTGTATACCGTCGAGCCGCGCAATGAGCATGATGGCTTCA 1085
    :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
348  AlValAlaMet.....GlyAlaSerAsnThrSerMetAlaMetG 361
1086  GGTCTCTCTCAGCATTTGGGAGACTGAGCTTGGCCATGCTGATGCTCA 1135
    :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
361  nAla.....AlaAlaLeuGlySerSerIleGlyThrAlaVal 374
1136  CGTGGTCTCTCTGACAGTTTGGCTGCTTAATCTTAATATGGAAGCA 1185
    :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
374  AlGlyGlyLeuMetThrGlyLeuIleLeuLysLeuProLeuThrGlyGln 390
1186  CCTCATGAGGCTTAATATTTTGTATGATGACCAAGTTTCTGGAAGTTCC 1233
    :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
391  ProSerAspGlnAsnCysTyrAspAspSerValTyrTrpLysValPro 406

seq_name: SwissProt_39:SRCH_HUMAN

seq_documentation_block:
ID  SRCH_HUMAN  STANDARD;  PRT;  699 AA.
AC  P23327;
DT  01-NOV-1991 (Rel. 20, Created)
DT  01-NOV-1991 (Rel. 20, Last sequence update)
DT  01-OCT-2000 (Rel. 40, Last annotation update)
DE  SARCOPLASMIC RETICULUM HISTIDINE-RICH CALCIUM-BINDING PROTEIN
DE  PRECURSOR.
GN  HRC OR HCP.
OS  Homo sapiens (human).
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX  NCBI_TaxID=9606;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  TISSUE=Skeletal muscle;
RX  MEDLINE=91244309; Pubmed=2037293;
RA  Hofmann S.L., Topham M., Hsieh C.L., Francke U.;
RT  cDNA and genomic cloning of HRC, a human sarcoplasmic reticulum
RT  protein, and localization of the gene to human chromosome 19 and
RT  mouse chromosome 7.
RL  Genomics 9:656-669(1991).
CC  -!- FUNCTION: HCP MAY PLAY A ROLE IN THE REGULATION OF CA(2+)
CC  SEQUESTRATION OR RELEASE IN THE SR OF SKELETAL AND CARDIAC
CC  MUSCLE.
CC  -!- SUBCELLULAR LOCATION: SARCOPLASMIC RETICULUM LUMEN.
CC  -!- SIMILARITY: STRONG, TO RABBIT HRC.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----

```

DR EMBL; M60052; AAA88071.1; -
 KM MIM; 142705; -
 FT CHAIN 1 28 SARCOPHAGIC RETICULUM HISTIDINE-RICH
 FT SIGNAL 29 699 CALCIUM-BINDING PROTEIN.
 FT DOMAIN 193 204 GLU-RICH (ACIDIC).
 FT DOMAIN 246 261 ASP-RICH (ACIDIC).
 FT DOMAIN 106 365 4 X TANDEM REPEATS, ACIDIC.
 FT REPEAT 180 213 1-1.
 FT REPEAT 238 270 1-2.
 FT REPEAT 295 318 1-3.
 FT REPEAT 343 365 1-4.
 FT DOMAIN 106 342 6 X APPROXIMATE TANDEM REPEATS.
 FT REPEAT 106 121 2-1.
 FT REPEAT 134 154 2-2.
 FT REPEAT 155 177 2-3.
 FT REPEAT 214 237 2-4.
 FT REPEAT 271 294 2-5.
 FT REPEAT 319 342 2-6.
 FT DOMAIN 627 673 METAL-BINDING (POTENTIAL).
 FT VARIANT 96 S -> V
 SQ SEQUENCE 699 AA; 80244 MW; 9922EEDF012C61DD CRC64;

alignment_scores:
 Quality: 141.00 Length: 407
 Ratio: 0.870 Gaps: 21
 Percent Similarity: 39.803 Percent Identity: 21.376

alignment_block:

US-09-600-714-41/rev x SRCH_HUMAN

Align seg 1/1 to: SRCH_HUMAN from: 1 to: 699

1131 AGCATATACAGATGCCAAGCTGAGTCCCAATGCTGAGAGAGACTG.. 1084
 ||||| ||| ||| |||||
 84 SerHisProAspArgGluLysGlu.....AspGluAspValSer 96
 1083GAAGCCATCATGTCATGCGCGCTCGAGCGGTATCAAGCCAGC 1038
 ||| ||| ||||| ||| ||| |||
 96 rLysGluTyrGluHisLeuLeuProGluHis.....ArgSerGln 110
 1037 AGCAATATGATGATCTCTCCAGACAGACCCAGCAAGTGAATGTA 988
 ||||| ||| ||| ||| |||||
 110 SPHisLysValGlyAsp.....GluGlyValSerGlyGluGluValPhe 124
 987 GCCCATGATGA.....GCTGTGGGGAATCCCGACACTCGGTACAC 944
 ||||| ||| ||| ||| |||
 125 AlaGluHisGlyGlyGlnAlaArgGlyHisArgGlyHis..... 137
 943 ACCCGGAGGACTTGCTCCCGCGAGGAGATCAG..... 907
 ||||| ||| ||| |||||
 138GlySerGluAspThrGluAspSerAlaGluHis 149
 906CCAGCCCAAGACCCAGACAGTGGCAAGCCAGCGAGA 868
 ||||| ||| ||| ||| |||||
 149 rGHisHisLeuProSerHisArgSerHisSerHis..GlnAspGluAsp 164
 867 AGGATCAGGTGACACAGAGTACCCAGACAGCGCTCTGCAACACCG 818
 ||||| ||| ||| ||| |||
 165 GluAspGluValAlaSerSerGluHisHis..... 175
 817 CACTGTGCATTAAGTCTTGCTGATCTTCCTTGGGGGTAGCCAGGAT 768
 |||||
 176HisLeu..... 177
 767 GACCTGATGATGGCTGTCCACACGCTGACTCTACAGCATAGTGTGT 718
 |||||
 178Leu 178
 717 GAACAGCGCATTTCTTCTTGATGAGCTTCTCAGACAGACAGATTGA 668

179 ArgHisGlyHisArgGlyHisAspGlyGluAspArgGlyGlyGluGlu 195
 :||| ||| :|||
 667 AACTTGGCCACACATCCCAAGAAAGAGCGGCC..CAGATGGCAGAC 621
 ||||| ||||| ||||| |||||
 195 uGluGluGluGluGluGluGluGluGluAlaSerThrGlyTyrGlyHisG 212
 620 AAACGTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 571
 ||||| ||||| ||||| |||||
 212 Lh..... 212
 570 AGGCTTGGACGACGACGACGACGACGACGACGACGACGACGACGAC 521
 ||||| ||||| ||||| |||||
 213AlaHisArgHisArgGlyHisGlySerGlu 223
 520 CGTATGATGTCATGATGTCATGTCATGTCATGTCATGTCATGTCATG 471
 ||||| ||||| ||||| |||||
 223AspGluAspValSerA 229
 470 ATGACATCTCTCAGGTTGCTTAAGCTGTCACTCCACACGAC..... 427
 ||||| ||||| ||||| |||||
 229 spGlyHis.....HisHisGlyPro 236
 426 ..CATCACCACCACTGCGCCCAAGTTGACCTCCCAAGACACATCCA 380
 ||||| ||||| ||||| |||||
 237 SerHisArgHisGlyHisGlyHisGlyHisGlyHisGlyHisGlyHis 253
 379 CTGAGATCAG..... 370
 253 pAspAspAspAspAspAspAspAspValSerIleGlyTyrArgHisGln 270
 ||||| ||||| ||||| |||||
 369 ..CACCCCAAGACACATCATGCTGCGCCAGCCCAATGATGACATGAT 322
 ||||| ||||| ||||| |||||
 270 LhHisArgHisGlyHisGlyHisGlyHisGlyHisGlyHisGlyHisGly 286
 321 GAACACCTTCCCAAGAGGAACTGCTCAGAGACCGTCCACGACGATG 272
 ||||| ||||| ||||| |||||
 287 GlyHisHisHisArgAsp..... 292
 271 CCCACGTCACACCAAGCGCCAGCATGAAGTTGAA..... 235
 ||||| ||||| ||||| |||||
 293 .ProSerHisArgHisArgSerHisGlyHisGlyHisGlyHisGlyHis 309
 234GCGCCACACTGCTCCAGCTGTGTCTCCGAACT 202
 ||||| ||||| ||||| |||||
 309 spValSerThrGlyTyrGlyHisGlnAlaHisArgHisGlnAspHisArg 325
 201 CGAGTGAAGCAAGCCCAAGCAATGCGCGCCATCAGGTCAGATCTTGGC 152
 ||||| ||||| ||||| |||||
 326 LysGluGluValGluAlaValSerGlyGlyHisHisHisHisValProAs 342
 151 CAACCTGATAGATGCCACGACGCGCTTTGATGCTCTTAAGAGCGCTCA 102
 ||||| ||||| ||||| |||||
 342 PHisArg.....HisGln..... 346
 101 TAGTGGTAAATAAATAAGAGAGAGATGAGAGCTCTCCAGTGTAG 52
 ||||| ||||| ||||| |||||
 347GlyHisArgAspGluGluGluAspGluAspValSerThrGluAr 361
 51 GCGCCAGAGGCGGCGAGCGAG 33
 ||||| ||||| ||||| |||||
 361 gTyrHisGlnGlyProGln 367
 seq_name: SwissProt_39:Y108_SYNY3
 seq_documentation_block:
 ID Y108_SYNY3 STANDARD; PRT; 507 AA.
 AC P54147;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)
 DE PUTATIVE AMMONIUM TRANSPORTER SL0108.
 GN SL0108.

OS Synechocystis sp. (strain PCC 6803).
 OC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
 OX NCBI_TaxID=1148;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=66127529; PubMed=8590279;
 RA Kaneko T., Tanaka A., Sato S., Kotani H., Sazuka T., Miyajima N.,
 Sugita M., Tabata S.;
 RT "Sequence analysis of the genome of the unicellular cyanobacterium
 Synechocystis sp. strain PCC6803. I. Sequence features in the 1 Mb
 region from map positions 64k to 92k of the genome.";
 RL DNA Res. 2:153-166(1995).
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (PROBABLE).
 CC -1- SIMILARITY: BELONGS TO THE AMT/MEP/NRGA FAMILY OF AMMONIUM
 CC TRANSPORTERS (TC 2.49).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL, D64004; BAA10631.1; -;
 DR InterPro: IPR001905; -;
 DR Pfam: PF00909; Ammonium_transp; 1.
 DR PROSITE: PS01219; AMMONIUM_TRANS; 1.
 KW Hypothetical protein; Transport; Transmembrane.
 FT TRANSMEM 44 64 POTENTIAL.
 FT TRANSMEM 90 110 POTENTIAL.
 FT TRANSMEM 128 148 POTENTIAL.
 FT TRANSMEM 154 174 POTENTIAL.
 FT TRANSMEM 178 198 POTENTIAL.
 FT TRANSMEM 211 231 POTENTIAL.
 FT TRANSMEM 245 265 POTENTIAL.
 FT TRANSMEM 286 306 POTENTIAL.
 FT TRANSMEM 315 335 POTENTIAL.
 FT TRANSMEM 347 367 POTENTIAL.
 FT TRANSMEM 370 390 POTENTIAL.
 FT TRANSMEM 399 419 POTENTIAL.
 FT TRANSMEM 438 458 POTENTIAL.
 SQ SEQUENCE 507 AA; 53580 MW; 1840AE99FFA08BE6A CRC64;

alignment_scores:
 Quality: 138.50 Length: 369
 Ratio: 0.710 Gaps: 18
 Percent Similarity: 52.846 Percent Identity: 23.306

alignment_block:
 US-09-600-714-41 x Y108_SYNY3 ..

Align seg 1/1 to: Y108_SYNY3 from: 1 to: 507

```

19 CGGTCGTCCGGCGC.....TGCCTGCCCT 44
   ||| |||||
34 ArgpHeValArgylsIleAsnSerProtrpLeuAlaCysValProle 50
   ||| |||||
45 CTGGGCGCTTAACACTG.....GAAGCAGCTCTCATTTCTCTCTCT 85
   ||| |||||
50 uThrAlaLeuIleValAlaIleTrpAsnAlaAlaIleAlaIleAspT 67
   ||| |||||
86 ATTTTATACCATATAGCGCTTCTAGAGATCAAAAGGCGCTG 135
   ||| |||||
67 hrgIuIleVal.....AsnIleThrValGluThrValAsnGluAsnVal 81
   ||| |||||
136 GCATCTATCA.....GTGGCCAGA 158
   ||| |||||
82 AlaThrLeuGlnGlyThrLeuAsnAlaIleTrpIleLeuIleAlaIle 98
   ||| |||||
159 TCTGACCGTATGGCGCATTTGGCTTCCCTCACTGAGATTTC 207
   ||| |||||

```

```

98 eLeuValIlePheMetAsnAlaGlyPheGlyMetLeuGluThrGlyLeuC 115
208 .....CGAGACACAGCTGAGACAGTGGCTTCAACTTCATGCTG 252
   ||| |||||
115 ysArgGlnIlyAsnAlaValAsnIleLeuThrIlyAsnLeuIleValPhe 131
   ||| |||||
253 CGCGTTGCT.....GTGCAGTGGCAATCTCTGTCAGCGCTTC 291
   ||| |||||
132 AlaLeuAlaThrIleAlaIleTrpAlaIle.....GlyPheSerIle 145
291 .....
145 uMetPheGlySerSerGlyAsnProPheValGlyPheGlyIlyPhePhe 162
292 .....CTGACCGATTCCTCTTGGAG 315
   ||| |||||
162 euserGlyAspHisThrAsnIlyGlyLeuSerProPheProGlnGlyLeu 178
316 GTGGTCATCACACTGTTAGTATTCGGCTGCCACACATGACTGTTGTC 365
   ||| |||||
179 ProValAlaValIlePheLeuPheGlnValAla...PheSerAlaThrAl 194
366 GGTGCTGATCTCAGTGCATGCTGCTGGAGAGTCAACTGGCGCAGT 415
   ||| |||||
194 AlaThrIleValSerGlyAlaValAlaGluArgIleIlyPheAsnGluP 211
416 TGTTGCTGATGCTGCTGAGAGTGCATGAGCTTACGCCACTGAGATG 465
   ||| |||||
211 helerIlePheSerValLeu.....LeuValGlyIleAlaIly 223
466 GTCATCATATATCTTCAACACAGC.....TACCACAT 500
   ||| |||||
224 ProIleThrGlyHisThrValIlyTrpAspAlaGlyIlyTrpLeuIlyTh 240
501 GAACATGATGCAC.....ATCACGTTGGAGGCT 532
   ||| |||||
240 GelyPheMetAspPheAlaGlySerThrValValHisSerValGlyGlyT 257
533 ATTTTGGCTGTCTGCTGCTGCTGCTGCCAAGCTCTACCCGAGGA 582
   ||| |||||
257 rPAlaAlaLeuAlaGlyAlaPheLeuLeuGlyProIlyGlyLeuIlyShe 273
583 ACGAGATTAAGATCAGACAGCAAGATPCCCACTTGTCT..... 624
   ||| |||||
274 ValAspGlyArg.....ProGlyAlaIleProGlyHisAsnMetGlyPh 288
625 .GCCATGCTGGCGGCTCTTCTGTGATGTTCTGGCCAGTTCACT 673
   ||| |||||
288 eAlaMetLeuGlyCysLeuIleLeuTrpIleGlyTrpPheGlyPheAsn 305
674 CTGCTCTGCTGAGAGTCCAAATCGAAGAAGAAATGCCGTTCACAC 723
   ||| |||||
305 roGly.....SerGlnLeuAlaAlaAspGlnAlaCys.....Ala 316
724 TACTATGCTTAGACAGTACGCGTGTGACAGCCATCTGAGGTCATCTT 773
   ||| |||||
317 TyrIleAlaValThrThrAsnLeuAlaIleSerAlaGlyGlyLeuThrAl 333
774 GGCTCAC.....CCCCAAGGAGATACGCAAGACTTATGTC 811
   ||| |||||
333 aThrPheThrSerTrpLeuIlyAspGlyIlyProAspLeuThrMetValI 350
812 ACAGTGGGTTGTTGGCAGAGCGCGTGGTGGTACCTCGTGCAGCTG 861
   ||| |||||
350 leAsnGlyValLeuAlaGlyLeuValGlyIleThrAlaGlyCysAlaGly 366
862 ATCCCTCTCGTGGCTTGCATGCTGCTGCTGCTGCTGCTGCTGCTGCT 911
   ||| |||||
367 Val...SerTyrTrpGlySerValIleIleGlyIleAlaGlyIleLe 382
912 CTCGCTC 918
   ||| |||||
382 uValVal 384

```

```

seq_name: swissprot_39: FHUB_ECOLI
seq_documentation_block:
ID FHUB_ECOLI STANDARD: PRT: 660 AA.
AC P06972; P77372;
DT 01-APR-1988 (Rel. 07, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE FERRICHRONE TRANSPORT SYSTEM PERMEASE PROTEIN FHUB (FERRICHRONE UPTAKE
DE PROTEIN FHUB).
GN FHUB.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87014116; PubMed=3020380;
RA Koester W., Braun V.;
RT "Iron hydroxamate transport of Escherichia coli: nucleotide sequence
RT of the fhb gene and identification of the protein.";
RL Mol. Genet. 204:435-442(1986).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=K12 / W3110;
RA Blatner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).
RN [4]
RP SEQUENCE FROM N.A.
RA Schramm S., Duncan M., Allen E., Araujo R., Aparicio A., Chung E.,
RA Davis K., Federspiel N., Hyman R., Kalman S., Komp C., Kurdi O.,
RA Lashari D., Lew H., Lin D., Namath A., Oefner P., Roberts D.,
RA Davis R.W.;
RT Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE OF 1-23 FROM N.A.
RX MEDLINE=88038363; PubMed=2823072;
RA Burkhardt R., Braun V.;
RT "Nucleotide sequence of the fhuc and fhud genes involved in iron
RT ATP-binding proteins.";
RL Mol. Genet. 209:49-55(1987).
CC -1- FUNCTION: THIS IS ONE OF THE PROTEINS INVOLVED IN THE HIGH-
CC AFFINITY TRANSPORT OF IRON(III)-FERRICHRONE INTO THE E.COLI
CC CELL.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. INNER MEMBRANE.
CC -1- SIMILARITY: WITH INTEGRAL MEMBRANE COMPONENTS OF OTHER BINDING-
CC PROTEIN-DEPENDENT TRANSPORT SYSTEMS. BELONGS TO THE FECD
CC SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X04319; CAA27852.1; -.

```

```

DR EMBL: X04319; CAA27852.1; ALT_INIT.
DR EMBL: D26562; CAB20273.1; -.
DR EMBL: AE000124; AAC73264.1; -.
DR EMBL: U70214; AAB08583.1; -.
DR EMBL: X05810; CAA29236.1; -.
DR PIR: S07318; S07318.
DR PIR: S45222; S45222.
DR EcoGene: EG10303; fhub.
DR InterPro: IPR000522; -.
DR Pfam: PF01032; Feecd_family_2.
KW Transport; Iron transport; Transmembrane; Inner membrane.
FT TRANSMEM 5
FT TRANSMEM 25
FT TRANSMEM 63 83
FT TRANSMEM 93 113
FT TRANSMEM 118 138
FT TRANSMEM 147 167
FT TRANSMEM 197 217
FT TRANSMEM 240 260
FT TRANSMEM 277 297
FT TRANSMEM 303 323
FT TRANSMEM 348 368
FT TRANSMEM 391 411
FT TRANSMEM 424 444
FT TRANSMEM 447 467
FT TRANSMEM 479 499
FT TRANSMEM 528 548
FT TRANSMEM 567 587
FT TRANSMEM 607 627
FT TRANSMEM 635 655
FT TRANSMEM 655 675
FT TRANSMEM 685 705
FT TRANSMEM 715 735
FT TRANSMEM 745 765
FT TRANSMEM 775 795
FT TRANSMEM 805 825
FT TRANSMEM 835 855
FT TRANSMEM 865 885
FT TRANSMEM 895 915
FT TRANSMEM 925 945
FT TRANSMEM 955 975
FT TRANSMEM 985 1005
FT TRANSMEM 1015 1035
FT TRANSMEM 1045 1065
FT TRANSMEM 1075 1095
FT TRANSMEM 1105 1125
FT TRANSMEM 1135 1155
FT TRANSMEM 1165 1185
FT TRANSMEM 1195 1215
FT TRANSMEM 1225 1245
FT TRANSMEM 1255 1275
FT TRANSMEM 1285 1305
FT TRANSMEM 1315 1335
FT TRANSMEM 1345 1365
FT TRANSMEM 1375 1395
FT TRANSMEM 1405 1425
FT TRANSMEM 1435 1455
FT TRANSMEM 1465 1485
FT TRANSMEM 1495 1515
FT TRANSMEM 1525 1545
FT TRANSMEM 1555 1575
FT TRANSMEM 1585 1605
FT TRANSMEM 1615 1635
FT TRANSMEM 1645 1665
FT TRANSMEM 1675 1695
FT TRANSMEM 1705 1725
FT TRANSMEM 1735 1755
FT TRANSMEM 1765 1785
FT TRANSMEM 1795 1815
FT TRANSMEM 1825 1845
FT TRANSMEM 1855 1875
FT TRANSMEM 1885 1905
FT TRANSMEM 1915 1935
FT TRANSMEM 1945 1965
FT TRANSMEM 1975 1995
FT TRANSMEM 2005 2025
FT TRANSMEM 2035 2055
FT TRANSMEM 2065 2085
FT TRANSMEM 2095 2115
FT TRANSMEM 2125 2145
FT TRANSMEM 2155 2175
FT TRANSMEM 2185 2205
FT TRANSMEM 2215 2235
FT TRANSMEM 2245 2265
FT TRANSMEM 2275 2295
FT TRANSMEM 2305 2325
FT TRANSMEM 2335 2355
FT TRANSMEM 2365 2385
FT TRANSMEM 2395 2415
FT TRANSMEM 2425 2445
FT TRANSMEM 2455 2475
FT TRANSMEM 2485 2505
FT TRANSMEM 2515 2535
FT TRANSMEM 2545 2565
FT TRANSMEM 2575 2595
FT TRANSMEM 2605 2625
FT TRANSMEM 2635 2655
FT TRANSMEM 2665 2685
FT TRANSMEM 2695 2715
FT TRANSMEM 2725 2745
FT TRANSMEM 2755 2775
FT TRANSMEM 2785 2805
FT TRANSMEM 2815 2835
FT TRANSMEM 2845 2865
FT TRANSMEM 2875 2895
FT TRANSMEM 2905 2925
FT TRANSMEM 2935 2955
FT TRANSMEM 2965 2985
FT TRANSMEM 2995 3015
FT TRANSMEM 3025 3045
FT TRANSMEM 3055 3075
FT TRANSMEM 3085 3105
FT TRANSMEM 3115 3135
FT TRANSMEM 3145 3165
FT TRANSMEM 3175 3195
FT TRANSMEM 3205 3225
FT TRANSMEM 3235 3255
FT TRANSMEM 3265 3285
FT TRANSMEM 3295 3315
FT TRANSMEM 3325 3345
FT TRANSMEM 3355 3375
FT TRANSMEM 3385 3405
FT TRANSMEM 3415 3435
FT TRANSMEM 3445 3465
FT TRANSMEM 3475 3495
FT TRANSMEM 3505 3525
FT TRANSMEM 3535 3555
FT TRANSMEM 3565 3585
FT TRANSMEM 3595 3615
FT TRANSMEM 3625 3645
FT TRANSMEM 3655 3675
FT TRANSMEM 3685 3705
FT TRANSMEM 3715 3735
FT TRANSMEM 3745 3765
FT TRANSMEM 3775 3795
FT TRANSMEM 3805 3825
FT TRANSMEM 3835 3855
FT TRANSMEM 3865 3885
FT TRANSMEM 3895 3915
FT TRANSMEM 3925 3945
FT TRANSMEM 3955 3975
FT TRANSMEM 3985 4005
FT TRANSMEM 4015 4035
FT TRANSMEM 4045 4065
FT TRANSMEM 4075 4095
FT TRANSMEM 4105 4125
FT TRANSMEM 4135 4155
FT TRANSMEM 4165 4185
FT TRANSMEM 4195 4215
FT TRANSMEM 4225 4235
FT TRANSMEM 4245 4255
FT TRANSMEM 4265 4275
FT TRANSMEM 4285 4295
FT TRANSMEM 4305 4315
FT TRANSMEM 4325 4335
FT TRANSMEM 4345 4355
FT TRANSMEM 4365 4375
FT TRANSMEM 4385 4395
FT TRANSMEM 4405 4415
FT TRANSMEM 4425 4435
FT TRANSMEM 4445 4455
FT TRANSMEM 4465 4475
FT TRANSMEM 4485 4495
FT TRANSMEM 4505 4515
FT TRANSMEM 4525 4535
FT TRANSMEM 4545 4555
FT TRANSMEM 4565 4575
FT TRANSMEM 4585 4595
FT TRANSMEM 4605 4615
FT TRANSMEM 4625 4635
FT TRANSMEM 4645 4655
FT TRANSMEM 4665 4675
FT TRANSMEM 4685 4695
FT TRANSMEM 4705 4715
FT TRANSMEM 4725 4735
FT TRANSMEM 4745 4755
FT TRANSMEM 4765 4775
FT TRANSMEM 4785 4795
FT TRANSMEM 4805 4815
FT TRANSMEM 4825 4835
FT TRANSMEM 4845 4855
FT TRANSMEM 4865 4875
FT TRANSMEM 4885 4895
FT TRANSMEM 4905 4915
FT TRANSMEM 4925 4935
FT TRANSMEM 4945 4955
FT TRANSMEM 4965 4975
FT TRANSMEM 4985 4995
FT TRANSMEM 5005 5015
FT TRANSMEM 5025 5035
FT TRANSMEM 5045 5055
FT TRANSMEM 5065 5075
FT TRANSMEM 5085 5095
FT TRANSMEM 5105 5115
FT TRANSMEM 5125 5135
FT TRANSMEM 5145 5155
FT TRANSMEM 5165 5175
FT TRANSMEM 5185 5195
FT TRANSMEM 5205 5215
FT TRANSMEM 5225 5235
FT TRANSMEM 5245 5255
FT TRANSMEM 5265 5275
FT TRANSMEM 5285 5295
FT TRANSMEM 5305 5315
FT TRANSMEM 5325 5335
FT TRANSMEM 5345 5355
FT TRANSMEM 5365 5375
FT TRANSMEM 5385 5395
FT TRANSMEM 5405 5415
FT TRANSMEM 5425 5435
FT TRANSMEM 5445 5455
FT TRANSMEM 5465 5475
FT TRANSMEM 5485 5495
FT TRANSMEM 5505 5515
FT TRANSMEM 5525 5535
FT TRANSMEM 5545 5555
FT TRANSMEM 5565 5575
FT TRANSMEM 5585 5595
FT TRANSMEM 5605 5615
FT TRANSMEM 5625 5635
FT TRANSMEM 5645 5655
FT TRANSMEM 5665 5675
FT TRANSMEM 5685 5695
FT TRANSMEM 5705 5715
FT TRANSMEM 5725 5735
FT TRANSMEM 5745 5755
FT TRANSMEM 5765 5775
FT TRANSMEM 5785 5795
FT TRANSMEM 5805 5815
FT TRANSMEM 5825 5835
FT TRANSMEM 5845 5855
FT TRANSMEM 5865 5875
FT TRANSMEM 5885 5895
FT TRANSMEM 5905 5915
FT TRANSMEM 5925 5935
FT TRANSMEM 5945 5955
FT TRANSMEM 5965 5975
FT TRANSMEM 5985 5995
FT TRANSMEM 6005 6015
FT TRANSMEM 6025 6035
FT TRANSMEM 6045 6055
FT TRANSMEM 6065 6075
FT TRANSMEM 6085 6095
FT TRANSMEM 6105 6115
FT TRANSMEM 6125 6135
FT TRANSMEM 6145 6155
FT TRANSMEM 6165 6175
FT TRANSMEM 6185 6195
FT TRANSMEM 6205 6215
FT TRANSMEM 6225 6235
FT TRANSMEM 6245 6255
FT TRANSMEM 6265 6275
FT TRANSMEM 6285 6295
FT TRANSMEM 6305 6315
FT TRANSMEM 6325 6335
FT TRANSMEM 6345 6355
FT TRANSMEM 6365 6375
FT TRANSMEM 6385 6395
FT TRANSMEM 6405 6415
FT TRANSMEM 6425 6435
FT TRANSMEM 6445 6455
FT TRANSMEM 6465 6475
FT TRANSMEM 6485 6495
FT TRANSMEM 6505 6515
FT TRANSMEM 6525 6535
FT TRANSMEM 6545 6555
FT TRANSMEM 6565 6575
FT TRANSMEM 6585 6595
FT TRANSMEM 6605 6615
FT TRANSMEM 6625 6635
FT TRANSMEM 6645 6655
FT TRANSMEM 6665 6675
FT TRANSMEM 6685 6695
FT TRANSMEM 6705 6715
FT TRANSMEM 6725 6735
FT TRANSMEM 6745 6755
FT TRANSMEM 6765 6775
FT TRANSMEM 6785 6795
FT TRANSMEM 6805 6815
FT TRANSMEM 6825 6835
FT TRANSMEM 6845 6855
FT TRANSMEM 6865 6875
FT TRANSMEM 6885 6895
FT TRANSMEM 6905 6915
FT TRANSMEM 6925 6935
FT TRANSMEM 6945 6955
FT TRANSMEM 6965 6975
FT TRANSMEM 6985 6995
FT TRANSMEM 7005 7015
FT TRANSMEM 7025 7035
FT TRANSMEM 7045 7055
FT TRANSMEM 7065 7075
FT TRANSMEM 7085 7095
FT TRANSMEM 7105 7115
FT TRANSMEM 7125 7135
FT TRANSMEM 7145 7155
FT TRANSMEM 7165 7175
FT TRANSMEM 7185 7195
FT TRANSMEM 7205 7215
FT TRANSMEM 7225 7235
FT TRANSMEM 7245 7255
FT TRANSMEM 7265 7275
FT TRANSMEM 7285 7295
FT TRANSMEM 7305 7315
FT TRANSMEM 7325 7335
FT TRANSMEM 7345 7355
FT TRANSMEM 7365 7375
FT TRANSMEM 7385 7395
FT TRANSMEM 7405 7415
FT TRANSMEM 7425 7435
FT TRANSMEM 7445 7455
FT TRANSMEM 7465 7475
FT TRANSMEM 7485 7495
FT TRANSMEM 7505 7515
FT TRANSMEM 7525 7535
FT TRANSMEM 7545 7555
FT TRANSMEM 7565 7575
FT TRANSMEM 7585 7595
FT TRANSMEM 7605 7615
FT TRANSMEM 7625 7635
FT TRANSMEM 7645 7655
FT TRANSMEM 7665 7675
FT TRANSMEM 7685 7695
FT TRANSMEM 7705 7715
FT TRANSMEM 7725 7735
FT TRANSMEM 7745 7755
FT TRANSMEM 7765 7775
FT TRANSMEM 7785 7795
FT TRANSMEM 7805 7815
FT TRANSMEM 7825 7835
FT TRANSMEM 7845 7855
FT TRANSMEM 7865 7875
FT TRANSMEM 7885 7895
FT TRANSMEM 7905 7915
FT TRANSMEM 7925 7935
FT TRANSMEM 7945 7955
FT TRANSMEM 7965 7975
FT TRANSMEM 7985 7995
FT TRANSMEM 8005 8015
FT TRANSMEM 8025 8035
FT TRANSMEM 8045 8055
FT TRANSMEM 8065 8075
FT TRANSMEM 8085 8095
FT TRANSMEM 8105 8115
FT TRANSMEM 8125 8135
FT TRANSMEM 8145 8155
FT TRANSMEM 8165 8175
FT TRANSMEM 8185 8195
FT TRANSMEM 8205 8215
FT TRANSMEM 8225 8235
FT TRANSMEM 8245 8255
FT TRANSMEM 8265 8275
FT TRANSMEM 8285 8295
FT TRANSMEM 8305 8315
FT TRANSMEM 8325 8335
FT TRANSMEM 8345 8355
FT TRANSMEM 8365 8375
FT TRANSMEM 8385 8395
FT TRANSMEM 8405 8415
FT TRANSMEM 8425 8435
FT TRANSMEM 8445 8455
FT TRANSMEM 8465 8475
FT TRANSMEM 8485 8495
FT TRANSMEM 8505 8515
FT TRANSMEM 8525 8535
FT TRANSMEM 8545 8555
FT TRANSMEM 8565 8575
FT TRANSMEM 8585 8595
FT TRANSMEM 8605 8615
FT TRANSMEM 8625 8635
FT TRANSMEM 8645 8655
FT TRANSMEM 8665 8675
FT TRANSMEM 8685 8695
FT TRANSMEM 8705 8715
FT TRANSMEM 8725 8735
FT TRANSMEM 8745 8755
FT TRANSMEM 8765 8775
FT TRANSMEM 8785 8795
FT TRANSMEM 8805 8815
FT TRANSMEM 8825 8835
FT TRANSMEM 8845 8855
FT TRANSMEM 8865 8875
FT TRANSMEM 8885 8895
FT TRANSMEM 8905 8915
FT TRANSMEM 8925 8935
FT TRANSMEM 8945 8955
FT TRANSMEM 8965 8975
FT TRANSMEM 8985 8995
FT TRANSMEM 9005 9015
FT TRANSMEM 9025 9035
FT TRANSMEM 9045 9055
FT TRANSMEM 9065 9075
FT TRANSMEM 9085 9095
FT TRANSMEM 9105 9115
FT TRANSMEM 9125 9135
FT TRANSMEM 9145 9155
FT TRANSMEM 9165 9175
FT TRANSMEM 9185 9195
FT TRANSMEM 9205 9215
FT TRANSMEM 9225 9235
FT TRANSMEM 9245 9255
FT TRANSMEM 9265 9275
FT TRANSMEM 9285 9295
FT TRANSMEM 9305 9315
FT TRANSMEM 9325 9335
FT TRANSMEM 9345 9355
FT TRANSMEM 9365 9375
FT TRANSMEM 9385 9395
FT TRANSMEM 9405 9415
FT TRANSMEM 9425 9435
FT TRANSMEM 9445 9455
FT TRANSMEM 9465 9475
FT TRANSMEM 9485 9495
FT TRANSMEM 9505 9515
FT TRANSMEM 9525 9535
FT TRANSMEM 9545 9555
FT TRANSMEM 9565 9575
FT TRANSMEM 9585 9595
FT TRANSMEM 9605 9615
FT TRANSMEM 9625 9635
FT TRANSMEM 9645 9655
FT TRANSMEM 9665 9675
FT TRANSMEM 9685 9695
FT TRANSMEM 9705 9715
FT TRANSMEM 9725 9735
FT TRANSMEM 9745 9755
FT TRANSMEM 9765 9775
FT TRANSMEM 9785 9795
FT TRANSMEM 9805 9815
FT TRANSMEM 9825 9835
FT TRANSMEM 9845 9855
FT TRANSMEM 9865 9875
FT TRANSMEM 9885 9895
FT TRANSMEM 9905 9915
FT TRANSMEM 9925 9935
FT TRANSMEM 9945 9955
FT TRANSMEM 9965 9975
FT TRANSMEM 9985 9995
FT TRANSMEM 10005 10015
FT TRANSMEM 10025 10035
FT TRANSMEM 10045 10055
FT TRANSMEM 10065 10075
FT TRANSMEM 10085 10095
FT TRANSMEM 10105 10115
FT TRANSMEM 10125 10135
FT TRANSMEM 10145 10155
FT TRANSMEM 10165 10175
FT TRANSMEM 10185 10195
FT TRANSMEM 10205 10215
FT TRANSMEM 10225 10235
FT TRANSMEM 10245 10255
FT TRANSMEM 10265 10275
FT TRANSMEM 10285 10295
FT TRANSMEM 10305 10315
FT TRANSMEM 10325 10335
FT TRANSMEM 10345 10355
FT TRANSMEM 10365 10375
FT TRANSMEM 10385 10395
FT TRANSMEM 10405 10415
FT TRANSMEM 10425 10435
FT TRANSMEM 10445 10455
FT TRANSMEM 10465 10475
FT TRANSMEM 10485 10495
FT TRANSMEM 10505 10515
FT TRANSMEM 10525 10535
FT TRANSMEM 10545 10555
FT TRANSMEM 10565 10575
FT TRANSMEM 10585 10595
FT TRANSMEM 10605 10615
FT TRANSMEM 10625 10635
FT TRANSMEM 10645 10655
FT TRANSMEM 10665 10675
FT TRANSMEM 10685 10695
FT TRANSMEM 10705 10715
FT TRANSMEM 10725 10735
FT TRANSMEM 10745 10755
FT TRANSMEM 10765 10775
FT TRANSMEM 10785 10795
FT TRANSMEM 10805 10815
FT TRANSMEM 10825 10835
FT TRANSMEM 10845 10855
FT TRANSMEM 10865 10875
FT TRANSMEM 10885 10895
FT TRANSMEM 10905 10915
FT TRANSMEM 10925 10935
FT TRANSMEM 10945 10955
FT TRANSMEM 10965 10975
FT TRANSMEM 10985 10995
FT TRANSMEM 11005 11015
FT TRANSMEM 11025 11035
FT TRANSMEM 11045 11055
FT TRANSMEM 11065 11075
FT TRANSMEM 11085 11095
FT TRANSMEM 11105 11115
FT TRANSMEM 11125 11135
FT TRANSMEM 11145 11155
FT TRANSMEM 11165 11175
FT TRANSMEM 11185 11195
FT TRANSMEM 11205 11215
FT TRANSMEM 11225 11235
FT TRANSMEM 11245 11255
FT TRANSMEM 11265 11275
FT TRANSMEM 11285 11295
FT TRANSMEM 11305 11315
FT TRANSMEM 11325 11335
FT TRANSMEM 11345 11355
FT TRANSMEM 11365 11375
FT TRANSMEM 11385 11395
FT TRANSMEM 11405 11415
FT TRANSMEM 11425 11435
FT TRANSMEM 11445 11455
FT TRANSMEM 11465 11475
FT TRANSMEM 11485 11495
FT TRANSMEM 11505 11515
FT TRANSMEM 11525 11535
FT TRANSMEM 11545 11555
FT TRANSMEM 11565 11575
FT TRANSMEM 11585 11595
FT TRANSMEM 11605 11615
FT TRANSMEM 11625 11635
FT TRANSMEM 11645 11655
FT TRANSMEM 11665 11675
FT TRANSMEM 11685 11695
FT TRANSMEM 11705 11715
FT TRANSMEM 11725 11735
FT TRANSMEM 11745 11755
FT TRANSMEM 11765 11775
FT TRANSMEM 11785 11795
FT TRANSMEM 11805 11815
FT TRANSMEM 11825 11835
FT TRANSMEM 11845 11855
FT TRANSMEM 11865 11875
FT TRANSMEM 11885 11895
FT TRANSMEM 11905 11915
FT TRANSMEM 11925 11935
FT TRANSMEM 11945 11955
FT TRANSMEM 11965 11975
FT TRANSMEM 11985 11995
FT TRANSMEM 12005 12015
FT TRANSMEM 12025 12035
FT TRANSMEM 12045 12055
FT TRANSMEM 12065 12075
FT TRANSMEM 12085 12095
FT TRANSMEM 12105 12115
FT TRANSMEM 12125 12135
FT TRANSMEM 12145 12155
FT TRANSMEM 12165 12175
FT TRANSMEM 12185 12195
FT TRANSMEM 12205 12215
FT TRANSMEM 12225 12235
FT TRANSMEM 12245 12255
FT TRANSMEM 12265 12275
FT TRANSMEM 12285 12295
FT TRANSMEM 12305 12315
FT TRANSMEM 12325 12335
FT TRANSMEM 12345 12355
FT TRANSMEM 12365 12375
FT TRANSMEM 12385 12395
FT TRANSMEM 12405 12415
FT TRANSMEM 12425 12435
FT TRANSMEM 12445 12455
FT TRANSMEM 12465 12475
FT TRANSMEM 12485 12495
FT TRANSMEM 12505 12515
FT TRANSMEM 12525 12535
FT TRANSMEM 12545 12555
FT TRANSMEM 12565 12575
FT TRANSMEM 12585 12595
FT TRANSMEM 12605 12615
FT TRANSMEM 12625 12635
FT TRANSMEM 12645 12655
FT TRANSMEM 12665 12675
FT TRANSMEM 12685 12695
FT TRANSMEM 12705 12715
FT TRANSMEM 12725 12735
FT TRANSMEM 12745 12755
FT TRANSMEM 12765 12775
FT TRANSMEM 12785 12795
FT TRANSMEM 12805 12815
FT TRANSMEM 12825 12835
FT TRANSMEM 12845 12855
FT TRANSMEM 12865 12875
FT TRANSMEM 12885 12895
FT TRANSMEM 12905 12915
FT TRANSMEM 12925 12935
FT TRANSMEM 12945 12955
FT TRANSMEM 12965 12975
FT TRANSMEM 12985 12995
FT TRANSMEM 13005 13015
FT TRANSMEM 13025 13035
FT TRANSMEM 13045 13055
FT TRANSMEM 13065 13075
FT TRANSMEM 13085 13095
FT TRANSMEM 13105 13115
FT TRANSMEM 13125 13135
FT TRANSMEM 13145 13155
FT TRANSMEM 13165 13175
FT TRANSMEM 13185 13195
FT TRANSMEM 13205 13215
FT TRANSMEM 13225 13235
FT TRANSMEM 13245 13255
FT TRANSMEM 13265 13275
FT TRANSMEM 13285 13295
FT TRANSMEM 13305 13315
FT TRANSMEM 13325 13335
FT TRANSMEM 13345 13355
FT TRANSMEM 13365 13375
FT TRANSMEM 13385 13395
FT TRANSMEM 13405 13415
FT TRANSMEM 13425 13435
FT TRANSMEM 13445 13455
FT TRANSMEM 13465 13475
FT TRANSMEM 13485 13495
FT TRANSMEM 13505 13515
FT TRANSMEM 13525 13535
FT TRANSMEM 13545 13555
FT TRANSMEM 13565 13575
FT TRANSMEM 13585 13595
FT TRANSMEM 13605 13615
FT TRANSMEM 13625 13635
FT TRANSMEM 13645 13655
FT TRANSMEM 13665 13675
FT TRANSMEM 13685 13695
FT TRANSMEM 13705 13715
FT TRANSMEM 13725 13735
FT TRANSMEM 13745 13755
FT TRANSMEM 13765 13775
FT TRANSMEM 13785 13795
FT TRANSMEM 13805 13815
FT TRANSMEM 13825 13835
FT TRANSMEM 13845 13855
FT TRANSMEM 13865 13875
FT TRANSMEM 13885 13895
FT TRANSMEM 13905 13915
FT TRANSMEM 13925 13935
FT TRANSMEM 13945 13955
FT TRANSMEM 13965 13975
FT TRANSMEM 13985 13995
FT TRANSMEM 14005 14015
FT TRANSMEM 14025 14035
FT TRANSMEM 14045 14055
FT TRANSMEM 14065 14075
FT TRANSMEM 14085 14095
FT TRANSMEM 14105 14115
FT TRANSMEM 14125 14135
FT TRANSMEM 14145 14155
FT TRANSMEM 14165 14175
FT TRANSMEM 14185 14195
FT TRANSMEM 14205 14215
FT TRANSMEM 14225 14235
FT TRANSMEM 14245 14255
FT TRANSMEM 14265 14275
FT TRANSMEM 14285 14295
FT TRANSMEM 14305 14315
FT TRANSMEM 14325 14335
FT TRANSMEM 14345 14355
FT TRANSMEM 14365 14375
FT TRANSMEM 14385 14395
FT TRANSMEM 14405 14415
FT TRANSMEM 14425 14435
FT TRANSMEM 14445 14455
FT TRANSMEM 14465 14475
FT TRANSMEM 14485 14495
FT TRANSMEM 14505 14515
FT TRANSMEM 14525 14535
FT TRANSMEM 14545 14555
FT TRANSMEM 14565 14575
FT TRANSMEM 14585 14595
FT TRANSMEM 14605 14615
FT TRANSMEM 14625 14635
FT TRANSMEM 14645 14655
FT TRANSMEM 14665 14675
FT TRANSMEM 14685 14695
FT TRANSMEM 14705 14715
FT TRANSMEM 14725 14735
FT TRANSMEM 14745 14755
FT TRANSMEM 14765 14775
FT TRANSMEM 14785 14795
FT TRANSMEM 14805 14815
FT TRANSMEM 14825 14835
FT TRANSMEM 14845 14855
FT TRANSMEM 14865 14875
FT TRANSMEM 14885 14895
FT TRANSMEM 14905 14915
FT TRANSMEM 14925 14935
FT TRANSMEM 14945 14955
FT TRANSMEM 14965 14975
FT TRANSMEM 14985 14995
FT TRANSMEM 15005 15015
FT TRANSMEM 15025 15035
FT TRANSMEM 15045 15055
FT TRANSMEM 15065 15075
FT TRANSMEM 15085 15095
FT TRANSMEM 15105 15115
FT TRANSMEM 15125 15135
FT TRANSMEM 15145 15155
FT TRANSMEM 15165 15175
FT TRANSMEM 15185 15195
FT TRANSMEM 15205 15215
FT TRANSMEM 15225 15235
FT TRANSMEM 15245 15255
FT TRANSMEM 15265 15275
FT TRANSMEM 15285 15295
FT TRANSMEM 15305 15315
FT TRANSMEM 15325 15335
FT TRANSMEM 15345 15355
FT TRANSMEM 15365 15375
FT TRANSMEM 15385 15395
FT TRANSMEM 15405 15415
FT TRANSMEM 15425 15435
FT TRANSMEM 15445 15455
FT TRANSMEM 15465 15475
FT TRANSMEM 15485 15495
FT TRANSMEM
```


[illegible]

```

seq_name: SwissProt_39:MOKD_SCHPO
1123 .....GTGATGACCTCTACGCTGCT..... 1143
462 eulenilletleMetilealalaIaglyArgIglyPheSerProHisArg 478
1144 ...CTCCCTACAGCTTTGCTCTAAAT 1167
479 MetIeuAlaIaglyMetAlaIeuser 487

seq_name: SwissProt_39:MOKD_SCHPO
seq_documentation_block:
ID MOKD_SCHPO STANDARD; PRT; 2358 AA.
AC Q9Y719; O94638;
DT 01-OCT-2000 (Rel. 40, Created)
DT 01-OCT-2000 (Rel. 40, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE CELL WALL ALPHA-1,3-GLUCAN SYNTHASE MOK13 (EC 2.4.1.183).
MOK13 OR SPBC16D10.05.
OS Schizosaccharomyces pombe (fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RA Katayama S., Dai H., Arellano M., Perez P., Toda T.;
RT "Fission yeast alpha-glucan synthase MOK1 localizes closely with actin
RT and play a role essential for cell morphogenesis and protein kinase C
RT function."
RT Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
[2]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RA Wood V., Rajandream M.A., Barrell B.G., Brown D., Churcher C.M.;
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
-1- CATALYTIC ACTIVITY: UDP-GLUCOSE + [ALPHA-D-GLUCOSYL-(1,3)](N) =
UDP + [ALPHA-D-GLUCOSYL-(1,3)](N+1).
-1- SIMILARITY: BELONGS TO THE GLUCOSYLTRANSFERASE FAMILY 1.
-----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC CC the European Bioinformatics Institute. There are no restrictions on its
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (see http://www.isb.ch/announce/).
CC CC or send an email to license@isb-sib.ch).
-----
DR EMBL: AB018382; BAA76559.1; -
DR EMBL: AL0355637; CAB38509.1; -
DR InterPro: IPR000461; -
DR Pfam: PF00128; alpha-amy1ase; 1.
KW Cell wall; Transferase; Glycosyltransferase.
FT CONFLICT 120 120 V -> VRRVMLCSTNRY (IN REF. 1).
FT SEQUENCE 2358 AA; 269192 MW; A175577C9D8AD731 CRC64;

alignment_scores:
Quality: 135.50 Length: 494
Ratio: 0.619 Gaps: 22
Percent Similarity: 44.332 Percent Identity: 20.445

align seg 1/1 to: MOKD_SCHPO from: 1 to: 2358
100 TATGAGCTTCCTTAGAGATCAAAAGGGGCGTGGCATCCTATCAAGT 149
||||| ||||||| |||:||||: |||:::
1872 TTTASP.....GluSplysLysPheIleasnThrGluInuIshPhele 1885

```

